

STAT 340/CS 437 PROBLEM SOLUTIONS.

1. 1.1 The following data gives the arrival times and the service times that each customer will require for the first 13 customers in a single server queue. On arrival the customer either enters service if the server is free or joins the waiting line. When the server completes work on a customer, the next one in the queue (i.e. the one that has been waiting the longest) enters service.

Arrival Times	12	31	63	95	99	154	198	221	304	346	411	455	537
Service Times	40	32	55	48	18	50	47	18	28	54	40	72	12

- (a) Determine the departure times of the 13 customers
- (b) Repeat (a) when there are two servers and a customer can be served by either one.
- (c) Repeat (a) under the new assumption that when the server completes a service, the next customer to enter service is the last to join the line-i.e. the one who has been waiting the least time.

Results: a) 52, 84, 139, 187, 205, 255, 302, 320, 348, 402, 451, 527, 549.

b) 52, 63, 118, 143, 136, 204, 245, 239, 332, 400, 451, 527, 549.

c) 52, 84, 139, 157, 207, 254, 272, 320, 348, 402, 451, 527, 549.

Matlab Code

```
% Matlab code for question ONE
% Data
arrival =[12 31 63 95 99 154 198 221 304 346 411 455 537];
service=[40 32 55 48 18 50 47 18 28 54 40 72 12];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
departure(1)=arrival(1)+service(1);
for i=2:13
    if arrival(i)<departure(i-1)
        \ departure(i)=departure(i-1)+service(i);
    else
        \ departure(i)=arrival(i)+service(i);
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
departure2(1)=arrival(1)+service(1);
departure2(2)=arrival(2)+service(2);
for i=3:13
    if arrival(i)<min(departure2(i-1),departure2(i-2))
        \ departure2(i)= min(departure2(i-1),departure2(i-2))+service(i);
    else
        \ departure2(i)=arrival(i)+service(i);
    end
end
```

end
end

2. 2.6 The continuous random variable X has probability density function given by

$$f(x) = cx, \quad 0 < x < 1$$

Find c and $P[X > 1/2]$.

Since f is a probability density function we need to find the constant c so that $\int_0^1 f(x)dx = 1$ or $\int_0^1 cx dx = 1$ giving $c = 2$. Then

$$P(X > \frac{1}{2}) = \int_{1/2}^1 f(x)dx = x^2 \Big|_{1/2}^1 = (1 - 1/4) = 3/4.$$

3. 2.7 If X and Y have joint probability density function given by

$$f(x, y) = 2 \exp\{-(x + 2y)\}, \quad 0 < x < \infty, 0 < y < \infty$$

find

$$P[X < Y].$$

The joint probability density function is

$$f(x, y) = 2e^{-(x+2y)}, \quad 0 < x < \infty \text{ and } 0 < y < \infty.$$

Therefore

$$\begin{aligned} P(X < Y) &= \int \int_{\{(x,y); x < y\}} f(x, y) dx dy \\ &= \int_0^\infty \left\{ \int_0^y f(x, y) dx \right\} dy \\ &= \frac{1}{3}. \end{aligned}$$

4. 2.15 An airplane needs at least half its engines to safely complete its mission. If each engine independently functions with probability p , for what values of p is a two engine plane safer than a three engine plane?

Let c be the number of engines required to safely complete the mission. Then when there are n engines, $c = 1$ when $n = 2$ and $c = 2$ when $n = 3$. The number of engines that operate properly is X , a binomial random variable with parameters n and p . Therefore the probability that the mission is successful is

$$\begin{aligned} P_2 &= P(X \geq 1) = \sum_{x=1}^2 \binom{2}{x} p^x (1-p)^{2-x} \text{ when } n = 2 \\ P_3 &= P(X \geq 2) = \sum_{x=2}^3 \binom{3}{x} p^x (1-p)^{3-x} \text{ when } n = 3 \end{aligned}$$

When is the two-engine plane better? When $P_2 > P_3$ or for values of p for which

$$\begin{aligned} 2p(1-p) + p^2 &\geq 3p^2(1-p) + (1-p)^3 \\ 1 + 2p^3 - p^2 - p &\geq 0 \end{aligned}$$

and this is true for all $0 \leq p \leq 1$.

5. 2.16 For a binomial(n, p) random variable X show that $p(i) = P[X = i]$ first increases and then decreases reaching its maximum value when

$$i = [p(n+1)], \text{ the largest integer less than or equal to } p(n+1).$$

If X has a binomial distribution so that $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$, we can obtain the ratio

$$\frac{P(X = i)}{P(X = i-1)} = \frac{(n-i+1)p}{i(1-p)}$$

Note that this ratio is greater than one if

$$(n-i+1)p > i(1-p)$$

or equivalently if

$$i < (n+1)p.$$

By a similar argument this ratio is smaller than one if

$$i > (n+1)p.$$

This means that the terms are increasing i.e. $P(X = i) > P(X = i-1)$ as long as $i < (n+1)p$ and subsequently the terms are decreasing. It follows that the maximum value of $P(X = i)$ is achieved when i is the largest integer less than or equal to $(n+1)p$ or $i = [(n+1)p]$.

6. 2.17 If X and Y are independent binomial random variables with parameters (n, p) and (m, p) respectively, argue without calculations that $X + Y$ is binomial $(m+n, p)$.

Let X be binomial (n, p) and Y be Binomial (m, p) . Then we can obtain X as the number of successes in n independent trials of an experiment in which the probability of a success on each trial is p . Similarly Y is the number of successes in m independent trials of an experiment with the same probability of success on each trial. Therefore $X + Y$ is the total number of successes on $n+m$ trials. It follows that $X + Y$ is Binomial $(n+m, p)$.

7. 2.19 If X is a Poisson(λ) random variable show that

$$(a) E(X) = \lambda$$

(b) $var(X) = \lambda$

The expected value of a Poisson random variable is

$$\begin{aligned}
 E(X) &= \sum_{x=0}^{\infty} xP(X=x) \\
 &= \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\
 &= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
 &= \lambda.
 \end{aligned}$$

Note also that

$$\begin{aligned}
 E(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x e^{-\lambda}}{x!} \\
 &= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} \\
 &= \lambda^2.
 \end{aligned}$$

It follows that

$$\begin{aligned}
 var(X) &= E(X^2) - \lambda^2 \\
 &= E(X(X-1)) + E(X) - \lambda^2 \\
 &= \lambda.
 \end{aligned}$$

8. 2.20 Let X and Y be independent Poisson random variables with parameters λ_1 and λ_2 respectively. Use question number 6 and a limit theorem for binomial random variables to argue heuristically that $X + Y$ has a $Poisson(\lambda_1 + \lambda_2)$ distribution. Then give an analytic proof of this result.

The heuristic argument is as follows. A binomial(n, p) random variable with a small value of p and large n is approximately $Poisson(\lambda)$ with $\lambda = np$. Therefore X is approximately Binomial(n_1, p) and Y Binomial(n_2, p) where $n_1 = \lambda_1/p$ and $n_2 = \lambda_2/p$ and p is chosen very small. By 2.17 it follows that the sum $X + Y$ is approximately binomial ($n_1 + n_2, p$) which is, since $n_1 + n_2$ is large and p is small is approximately $Poisson((n_1 + n_2)p)$ or $Poisson(\lambda_1 + \lambda_2)$. The formal argument is as follows. Suppose X has a $Poisson(\lambda_1)$ distribution

and Y has a $Poisson(\lambda_2)$ distribution and they are independent. Then

$$\begin{aligned}
 P[X + Y = k] &= \sum_{i=0}^k P(X = i, Y = k - i) \\
 &= \sum_{i=0}^k \frac{\lambda_1^i e^{-\lambda_1}}{i!} \frac{\lambda_2^{k-i} e^{-\lambda_2}}{(k-i)!} \\
 &= \frac{1}{k!} e^{-(\lambda_1 + \lambda_2)} \sum_{i=0}^k \binom{k}{i} \lambda_1^i \lambda_2^{k-i} \\
 &= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)}
 \end{aligned}$$

and this is the Poisson distribution with parameter $(\lambda_1 + \lambda_2)$. So the sum of independent Poisson random variables is also Poisson with the parameters added.

9. 2.28 If X is an exponential random variable with rate parameter λ show that

(a) $E(X) = \frac{1}{\lambda}$

(b) $var(X) = \frac{1}{\lambda^2}$.

For an exponential random variable X , with probability density function $f(x) = \lambda e^{-\lambda x}, x > 0$

$$\begin{aligned}
 E(X) &= \int_0^{\infty} x \lambda e^{-\lambda x} dx \\
 &= 1/\lambda.
 \end{aligned}$$

and

$$\begin{aligned}
 E(X^2) &= \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx \\
 &= 2/\lambda^2.
 \end{aligned}$$

Therefore the variance is

$$var(X) = E(X^2) - (EX)^2 = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2.$$

10. 2.29 Persons A , B , and C are waiting at a bank having two tellers when it opens in the morning. Persons A and B each go to a teller and C waits in line. If the times required to serve customers are independent exponential(λ) random variables, what is the probability that C is the last to leave the bank? (No computation is required)

Let us suppose that B leaves the bank after A . Then by the memoryless property of the Exponential distribution, at the moment that A leaves the bank, C has remaining service time which is exponentially distributed with parameter λ , and so does B . These two service times have the same continuous exponential distribution are independent. Therefore the probability that C 's service time is the larger is $\frac{1}{2}$.

11. 2.30 Let X and Y be independent exponential random variables with rate parameters λ and μ respectively. Show that

$$P[X < Y] = \frac{\lambda}{\lambda + \mu}.$$

The joint probability density function, since X and Y are independent exponential is

$$f(x, y) = \lambda\mu e^{-\lambda x} e^{-\mu y}, x > 0, y > 0.$$

Therefore

$$\begin{aligned} P(X < Y) &= \int \int_{\{(x,y); x < y\}} f(x, y) dx dy \\ &= \int_0^\infty \left\{ \int_0^y \lambda e^{-\lambda x} dx \right\} \mu e^{-\mu y} dy \\ &= \frac{\lambda}{\lambda + \mu}. \end{aligned}$$

12. 2.31 Consider a Poisson process in which events occur at a rate of 0.3 per hour. What is the probability that not events occur between 10 AM and 2 P.M.?

Let $N(t)$ be the number of events occurring in an interval of length t . According to the properties of a Poisson process, $N(t)$ has a Poisson distribution with parameter given by λt where λ is the intensity of the Poisson process. In this case the intensity is $\lambda = .3$ and so $N(4)$ has a Poisson distribution with parameter 1.2. Therefore

$$P[N(4) = 0] = e^{-1.2}$$

13. 2.32 For a Poisson process with rate λ , define $N(s)$ = the number of events in the interval $[0, s]$. Find

$$P[N(s) = k | N(t) = n]$$

for $s < t$.

For a Poisson process, recall that $N(t)$ is the number of events in the interval $[0, t]$. Therefore for $s < t$ and $0 \leq k \leq n$,

$$\begin{aligned} P[N(s) = k | N(t) = n] &= \frac{P[k \text{ events in } [0, s] \text{ and } n - k \text{ events in } [s, t]]}{P[n \text{ events in } [0, t]]} \\ &= \frac{n! (\lambda s)^k e^{-\lambda s} (\lambda(t-s))^{n-k} e^{-\lambda(t-s)}}{k! (n-k)! (\lambda t)^n e^{-\lambda t}} \\ &= \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}. \end{aligned}$$

The conditional distribution of the number of events in a subinterval given the number in an interval is binomial.

14. 2.36 If X and Y are independent exponential random variables, show that the conditional distribution of X given $X + Y = t$ is the uniform distribution on $(0, t)$.

Suppose that $f(x, y)$ denotes the joint probability density function of X, Y and f_X and f_T denote the probability density functions of X and $T = X + Y$ respectively. Then the conditional probability density function of X given $T = t$ is

$$f_{X|T}(x|t) = \frac{f_X(x)f_Y(t-x)}{f_T(t)} \text{ for } 0 \leq x \leq t.$$

In this case both X and Y are exponentially distributed with identical probability density functions $f_X(x) = \lambda e^{-\lambda x}, x > 0$. This is a special case of the gamma density function and the sum of independent gamma random variables with the same scale parameter is also gamma. In other words, T has the $\text{Gamma}(2, 1/\lambda)$ probability density function

$$f_T(t) = \lambda^2 t e^{-\lambda t}, t > 0$$

Substituting

$$\begin{aligned} f_{X|T}(x|t) &= \frac{f_X(x)f_Y(t-x)}{f_T(t)} \\ &= \frac{\lambda e^{-\lambda x} \lambda e^{-\lambda(t-x)}}{\lambda^2 t e^{-\lambda t}} \\ &= \frac{1}{t} \text{ for } 0 \leq x \leq t. \end{aligned}$$

In other words given that the sum of two independent exponential random variables is t the conditional distribution of one of them, say X , is uniform on the interval of possible values of X , namely $[0, t]$.

15. 3.1 the sequence is 5, 15, 45, 135, 105, 15, 45, 135, 105, 15, 45, Note that the sequence repeats after the four distinct numbers 15, 45, 135, 105 and so the sequence has period 4.

16. 3.2 If $x_0 = 3$ and

$$x_n = (5x_{n-1} + 7) \bmod 200$$

find x_1, \dots, x_{10} and the period of the generator.

Using $\text{lcg}(3, 5, 7, 200, 20)$ we obtain the first 20 values of x_n , 3 22 117 192 167 42 17 92 67 142 117 192, Notice that the values 3 and 22 never recur again but 117 appears again after 8 numbers. Indeed after 117 the sequence cycles through 117 192 167 42 17 92 67 142 repeatedly and the period is 8.

17. 3.3 Use simulation to approximate the integral

$$\int_0^1 \exp\{e^x\} dx$$

The integral cannot be obtained analytically. However a numerical approximation gives

$$\int_0^1 \exp(e^x) dx \approx 6.3166$$

We can approximate this integral using Monte Carlo with the following Matlab Code:

```
T=rand(1,100000); mean(exp(exp(T)))
```

18. Use the runs (up) test and the normal approximation to its distribution to determine whether the following sequence is significantly different from independent $U[0, 1]$ variables.

.427,.321,.343,.474,.558,.746,.032,.522,.604,.745,.251,.310,
.798,.037,.081,.95,.812,.453,.644,.920,.951,.146,.155,.429,.525,.2,.219,.675,.845,.
676

In this case $n = 30$ and the observed number of runs is $R = 15$. The expected number and standard deviation under the hypothesis that the numbers are independent Uniform[0,1] is

$$E(R) = \frac{2n-1}{3} = 19.67, SD(R) = \sqrt{\frac{16n-29}{90}} = 2.24$$

In this case, the observed value 15 is within $2SD(R)$ of the expected value $E(R)$ and so there is no significant evidence against the hypothesis of independent uniform. (note: this is a test of independence that can be applied for any continuous random variables- it is not sensitive to the particular continuous distribution)

19. Let x_1, \dots, x_{30} be the supposed $U[0, 1]$ values listed in problem 18 and suppose we plot the 15 points $\{(x_{2j-1}, x_{2j}), j = 1, 2, \dots, 15\}$. Divide the unit square is divided into four equal subsquares, and count the number of points in each of these four regions. Use the chi-squared test to determine whether the chi-squared statistic is significant. The number of point in each of the four quadrants is 3, 4, 4, 4 respectively. The expected number under the assumption that the points are uniform in the square is $\frac{17}{4} = 3.75$ for each of the four quadrants. Therefore the Chi-square statistic is

$$\begin{aligned} \chi^2 &= \frac{(3-3.75)^2}{3.75} + \frac{(4-3.75)^2}{3.75} + \frac{(4-3.75)^2}{3.75} + \frac{(4-3.75)^2}{3.75} \\ &= 0.2 \end{aligned}$$

This should be compared with a chi-squared distribution having $4 - 1 = 3$ degrees of freedom and for this distribution the critical value at level 5% is 7.81. Since the observed value is well below this critical value, there is no evidence according to this test that the sequence does not consist of independent uniform random variables.

20. 3.4 *Use simulation to approximate the integral*

$$\int_0^1 (1-x^2)^{3/2} dx$$

This is similar to question 17. In this case the integral

$$\int_0^1 (1-x^2)^{3/2} dx$$

can be obtained analytically and it is

$$\left[-\frac{x}{8}(2x^2-5)\sqrt{1-x^2} + \frac{3}{8}\sin^{-1}(x)\right]_0^1 = \frac{3}{16}\pi$$

or around 0.58905. We can approximate this integral using Monte Carlo with the following Matlab Code:

```
T=rand(1,100000);
mean((1-T.^2).^1.5);
```

21. 3.5 *Use simulation to approximate the integral*

$$\int_{-2}^2 \exp\{x+x^2\} dx$$

In this case the integral

$$\int_{-2}^2 e^{x+x^2} dx$$

can be obtained in terms of the normal cumulative distribution function or the “error function” and is around 93.163. We can write this integral as an expected value of the form

$$4E(e^{X+X^2})$$

where X has a uniform distribution on the interval $[-2, 2]$. Therefore, we can approximate this integral using Monte Carlo with the following Matlab Code: the last line obtains the standard error of the estimator

```
U=-2+4*rand(1,100000);
4*mean(exp(U+U.^2))
se= 4*sqrt(var(exp(U+U.^2))/100000)
```

22. 3.6 *Use simulation to approximate the integral*

$$\int_0^\infty x(1+x^2)^{-2} dx$$

Ross suggests changing variables so that the new variable ranges over the unit interval to

$$\begin{aligned} y &= 1/(x+1), \text{ and } x = \frac{1}{y} - 1 \\ dy &= -1/(x+1)^2 dx = -y^2 dx \end{aligned}$$

So the integral becomes, after a change of variables

$$\begin{aligned} & \int_0^1 \left(\frac{1}{y} - 1\right) \left(1 + \left(\frac{1}{y} - 1\right)^2\right)^{-2} y^{-2} dy \\ &= \int_0^1 \left(\frac{1}{y} - 1\right) (2y^2 - 2y + 1)^{-1} dy \end{aligned}$$

and the true value of this integral is $1/2$. We can approximate this integral using Monte Carlo with the following Matlab Code: the last line obtains the standard error of the estimator

```
U=rand(1,100000);
T=(1/U-1)./(2*U.^2 -2*U+1)
mean(T)
se= sqrt(var(T)/100000)
```

23. 3.7 Use simulation to approximate the integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

The exact value of the integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

is $\sqrt{\pi}$, as can be shown by conversion to polar coordinates. We may rewrite as

$$\int_{-\infty}^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx$$

and then change variables as in the problem above to

$$\begin{aligned} y &= 1/(x+1), \text{ and } x = \frac{1}{y} - 1 \\ dy &= -1/(x+1)^2 dx = -y^2 dx \end{aligned}$$

so the integral becomes

$$2 \int_0^1 \exp\left\{-\left(\frac{1}{y} - 1\right)^2\right\} y^{-2} dy.$$

which can be estimated using MATLAB

```
U=rand(1,100000);
F=exp(-(1-1./U).^2)./U.^2;
2*mean(F)
ans = 1.7737 which is in error by about .001.
```

24. 3.8 Use simulation to approximate the integral

$$\int_0^1 \int_0^1 e^{(x+y)^2} dy dx$$

First write this as $E(\exp(X + Y)^2)$ where X and Y are independent $U[0, 1]$.

Then in matlab:

```
x=(rand(1,100000)); y=(rand(1,100000));
f=exp((x+y).^2); m=mean(f)
ans=4.8924
```

25. 3.9 Use simulation to approximate the integral

$$\int_0^\infty \int_0^x e^{-(x+y)} dy dx$$

The true value of the integral

$$\int_0^\infty \int_0^x e^{-(x+y)} dy dx$$

is 1/2. Consider a change of variables

$$\begin{aligned} t &= 1/(x+1), & x &= 1/t - 1 \\ s &= y/x, & y &= sx = s\left(\frac{1}{t} - 1\right) \end{aligned}$$

so that now both variables are in the unit square. Then

$$\begin{aligned} \int_0^\infty \int_0^x e^{-(x+y)} dy dx &= \int_0^1 \int_0^1 \exp\left\{-\left(\frac{1}{t} - 1\right)(1 + s)\right\} \left| \frac{\partial x}{\partial t} \frac{\partial x}{\partial s} \right| ds dt \\ &= \int_0^1 \int_0^1 \exp\left\{-\left(\frac{1}{t} - 1\right)(1 + s)\right\} t^{-2} \left| \frac{1}{t} - 1 \right| ds dt \end{aligned}$$

This integral can be approximated using two independent uniform[0, 1], say S, T and is equal to

$$E\left[\exp\left\{-\left(\frac{1}{T} - 1\right)(1 + S)\right\} T^{-2} (T^{-1} - 1)\right].$$

or in matlab

```
S=rand(1,100000); T=rand(1,100000); mean(exp(-(1./T-1).*(1+S)).*(1./T.^2).*(1./T-1))
```

26. 3.10 Use simulation to approximate

$$\text{cov}(U, e^U)$$

where U is uniform[0, 1]. Compare your approximation with the exact value.

In this case we can write

$$\text{cov}(U, e^U) = E(e^U (U - \frac{1}{2}))$$

and so this gives the following matlab code

`U=rand(1,100000); mean(exp(U).*(U-.5))` providing a value of around 0.1405.
Alternatively use the sample covariance. In matlab if we use `cov(U,exp(U))` we obtain the sample covariance matrix

$$\begin{bmatrix} .0832 & .1406 \\ .1406 & .2416 \end{bmatrix}$$

and the covariance is the off-diagonal element 0.1406.

27. 3.11 Let U be uniform $[0, 1]$. Use simulation to approximate

(a) $\text{cov}(U, \sqrt{1-U^2})$

(b) $\text{cov}(U^2, \sqrt{1-U^2})$

The easiest approximation is to use the sample covariance.

`U=rand(1,100000);`

`V=sqrt(1-U.^2);`

`mean(U.*V)-mean(U)*mean(V)`. Gives value -0.0597.

Similarly

`mean((U.^2).*V)-mean(U.^2)*mean(V)` returns value -0.06578.

28. 3.12 Let U_1, U_2, \dots be uniform $[0, 1]$ random variables and define

$$N = \min\{n; \sum_{i=1}^n U_i > 1\}.$$

Estimate $E(N)$ by generating 100, 1000, 10,000 and 100,000 values. What is the true value of $E(N)$?

For the expected value of N , note that

$$P(N \geq 2) = 1$$

$$P(N \geq 3) = P(U_1 + U_2 \leq 1) = 1/2$$

$$P(N \geq 4) = P(U_1 + U_2 + U_3 \leq 1) = \frac{1}{3!}$$

$$P(N \geq j) = \frac{1}{(j-1)!}$$

Now notice that $\sum_{j=1}^{\infty} P(N \geq j)$ can be written as

$$\begin{array}{rcllclcl} P(N \geq 1) & = & P(N = 1) & + & P(N = 2) & + & P(N = 3) & + & P(N = 4) & + & \dots \\ P(N \geq 2) & & & & P(N = 2) & & + & P(N = 3) & + & P(N = 4) & + \dots \\ P(N \geq 3) & & & & & & P(N = 3) & & + & P(N = 4) & + \dots \\ P(N \geq 4) & & & & & & & & P(N = 4) & & + \dots \\ \text{SUM} = & & P(N = 1) & & + 2P(N = 2) & & + 3P(N = 3) & & + 4P(N = 4) & & + \dots \end{array}$$

This shows that $E(N) = \sum_{j=1}^{\infty} jP(N = j) = \sum_{j=1}^{\infty} P(N \geq j) = 1 + 1 + 1/2 + \dots = e$.

We may generate values of N using the following function. The function is not perfect however, since it does not allow the event $N > 100$, an event which has probability less than 10^{-157} .

```
function N=generateN(m)
N=[];
for i=1:m
V=cumsum(rand(1,100));
N=[N 1+sum(V<1)];
end
```

29. 3.13 Let U_1, U_2, \dots be uniform $[0, 1]$ random variables and define

$$N = \max\{n; \prod_{i=1}^n U_i > e^{-3}\}.$$

(a) Find $E(N)$ by simulation.

(b) Find $P[N = i]$ by simulation. What do you think is the distribution of N ?

First note that we can write N as

$$N = \max\{n; \sum_{i=1}^n \ln(U_i) > -3\}.$$

$$N = \max\{n; \sum_{i=1}^n -\ln(U_i) < 3\}.$$

Since the random variables $-\ln(U_i)$ all have an exponential(1) distribution this counts the total number of points in a Poisson process with intensity (1) that lie in the interval $[0, 3]$ and therefore this has the Poisson distribution with parameter 3.

$$P[N = i] = \frac{3^i}{i!} e^{-3}, i = 0, 1, \dots$$

and $E(N) = 3$. The following Matlab function can be used to generate k values of N .

```
• function N=q29(k)
• %generates a total of k random variables as in ques-
  tion 29
• N=[];
• for i=1:k
• x=cumprod(rand(1,10));
• while x(length(x))>exp(-3)
• x=[x x(length(x))*cumprod(rand(1,10))];
• end
• N=[N sum(x>exp(-3))];
```

- end

For example the statement `mean(q29(1000))` provides the average of 1000 simulated values of N and `mean(3==q29(1000))` provides the value 0.2260 which estimates the probability $P(N = 3)$ (the true value is about 0.2240).

30. 4.1 Write a program to generate n values of a random variable X with

$$P[X = i] = i/3, i = 1, 2$$

Generate $n = 100, 1,000$ and $10,000$ values of the random variable and compare the proportion of values of X equal to 1.

The vector of length N of suitable values of X can be generated using `X=1+(rand(1,N)>1/3);`

31. 4.2 Write a Matlab function which takes a vector of probabilities (p_1, \dots, p_n) as input and outputs a single value of a random variable having this probability (mass) function.

```
function X=simul(p);
X=sum(rand>[0 cumsum(p)]) ;
```

32. 4.3 Give an efficient algorithm to simulate the value of a random variable X such that

$$P[X = 1] = 0.3, P[X = 2] = 0.2, P[X = 3] = 0.35, P[X = 4] = 0.15$$

Generate $U = \text{Uniform}(0,1)$
 If $U < 0.35$, set $x=3$ and stop
 if $U < 0.65$, set $x=1$ and stop
 if $U < 0.85$ set $x=2$ and stop
 else set $x=4$

33. 4.4 A deck of 100 cards numbered $1, 2, \dots, 100$ is shuffled and then turned over one card at a time. Say a “hit” occurs whenever card i is the i th card to be turned over. Write a Matlab program to estimate the expected value and the variance of the total number of hits. Find the true value and compare with your estimates.

Consider the probability that the i 'th card is a “hit”. Since all 100 cards are possible in the i 'th location, the probability of a hit on card i is $1/100$. Therefore if $I_i = 1$ or 0 as the i 'th card is a hit or not, then the total number of hits is $\sum_{i=1}^{100} I_i$. The expected value is therefore $100E(I_i) = 10$. We may generate values of N with the following two functions.

```
function p=permuternd(n)
% generates a random permutation of the integers 1:n
p=1:n;
for k=n:-1:2
i=ceil(k*rand); q=[p(i) p(k)]; p(k)=q(1) ;p(i)=q(2);
end
```

```

function f=matchesrnd(n,m)
% f is the observed number of matches of n items over m
trials.
f=[];
for i=1:m
f=[f sum(permuternd(n)==(1:n))];
end
f=matchesrnd(100,1000);
mean(f) = 0.980000000000000
var(f) = 0.94254254254255 (the true values are both =1)

```

34. 4.5 Consider the following recursive algorithm for generating a random permutation Π_n of the integers $\{1, 2, \dots, n\}$. Begin with a random permutation, $n = 1$, namely 1. To generate a random permutation Π_{i+1} from a random permutation $\Pi_i = \{j_1, j_2, \dots, j_i\}$, first adjoin the value $i + 1$ to obtain $\{j_1, j_2, \dots, j_i, i + 1\}$ and then interchange the last component $i + 1$ with component J , randomly selected from the $i + 1$ values $\{j_1, j_2, \dots, j_i, i + 1\}$ (so that $i + 1$ can end up being interchanged with itself). Show that this algorithm produces a random permutation and write a Matlab function which implements this algorithm. Test your algorithm by repeatedly generating permutations Π_{10} .

Consider the following function

```

function v=perm(n)
v=1;
for i=2:n;
j=ceil(i*rand);      v=[v j];      v([j i])=[i v(j)];
end

```

Running this function 1000 times as with

```

x=[]
for i=1:1000; x=[x;perm(10)]; end

```

generates a matrix x of 1000 such permutations, each one occupying a row of the matrix. One can check that the individual columns, say column 3 has the correct frequency with `mean(x(:,3))==1`

35. 4.6 Using $n = 100$ random numbers explain how to find an approximation for

$$\sum_{i=1}^N e^{i/N}$$

where $N = 10,000$. Find the approximation and indicate if it is a good one.

We wish to approximate a sum of the form

$$\sum_{i=1}^N e^{i/N}$$

with $N = 10,000$. Note that if we replace the sum by an average

$$\frac{1}{N} \sum_{i=1}^N e^{i/N}$$

then this is approximately

$$E(e^U)$$

where U is $U[0, 1]$. This in turn can be approximated using the average of 100 such observations

$$\frac{1}{100} \sum_{i=1}^{100} e^{U_i}$$

or in MATLAB,

```
mean(exp(rand(1,100)))
```

This gives approximately 1.72 or around $e - 1$.

36. 4.7 A pair of dice are to be continually rolled until all the possible outcomes $2, 3, \dots, 12$ have occurred at least once. Develop a simulation study to estimate the expected number of dice rolls that are needed.

The following matlab function will generate the number of dice rolls needed. If we define this function and then enter `v=q36(1000)` then the vector `v` consists the results of 1000 simulations and `mean(v)` produces 59.08, an estimate of $E(N)$.

```
function v=q36(nsim)
%nsim=number of simulations. output v=vector of results of nsim simulations
v=[];
for i=1:nsim
y=sum(ceil(6*rand(2,12))); % initializes with 12 rolls
while (length(unique(y))<11)
x=sum(ceil(6*rand(1,2))); %generates sum of two rolls of the dice
y=[y x]; % augments the list of y values until unique(y)=11
end
v=[v length(y)];
end
```

37. 5.1 Give a method for generating a random variable having density function

$$f(x) = \frac{e^x}{e-1}, 0 < x < 1$$

Using 10000 simulated values from this distribution, estimate its mean and variance. Use inverse transform. In this case $F(x) = (e^x - 1)/(e - 1)$ and $F^{-1}(U) = \ln(1 + U(e - 1))$. `x=log(1+rand(1,10000)*(exp(1)-1)); mean(x)=0.58` and `var(x)=.0792`. The true values are $\int_0^1 x \frac{e^x}{e-1} dx = 0.58198$ and the variance is $\int_0^1 x^2 \frac{e^x}{e-1} dx - (0.58198)^2 := 0.079323$.

38. 5.2 Give a method for generating a random variable having density function

$$f(x) = \begin{cases} \frac{x-2}{2} & \text{if } 2 \leq x \leq 3 \\ 1 - \frac{x}{6} & \text{if } 3 \leq x \leq 6 \end{cases}$$

Using 10000 simulated values from this distribution, estimate its mean and variance. Apply the inverse transform method. In this case

$$F(x) = \begin{cases} \frac{(x-2)^2}{4} & \text{if } 2 \leq x \leq 3 \\ \frac{1}{4} + (x-3)\left(\frac{3}{4} - \frac{x}{12}\right) & \text{if } 3 < x < 6 \end{cases}$$

Inverting this c.d.f. gives

$$F^{-1}(U) = \begin{cases} 2 + 2\sqrt{U} & \text{if } 0 \leq U \leq 1/4 \\ 6 - 2\sqrt{3-3U} & \text{if } U > 1/4 \end{cases}$$

39. 5.4 Give a method for generating a Weibull random variable having cumulative distribution function

$$F(x) = 1 - \exp(-\alpha x^\beta), 0 < x < \infty$$

Using 10000 simulated values from this distribution, estimate its mean and variance when $\alpha = 1, \beta = 2$. Again inverse transform. $X = (-\ln(1 - U)/\alpha)^{1/\beta}$. The probability density function is

$$f(x) = F'(x) = \alpha x^{\beta-1} \beta e^{-\alpha x^\beta}, 0 < x < \infty$$

and so the true mean value in the case $\alpha = 1, \beta = 2$

$$2 \int_0^\infty x^2 e^{-x^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{\pi}}$$

Similarly

$$E(X^2) = 2 \int_0^\infty x^3 e^{-x^2} dx = 1$$

and the variance is $1 - \frac{\pi}{4}$.

40. 5.5 Give a method for generating a random variable having density function

$$f(x) = \begin{cases} e^{2x} & \text{if } -\infty \leq x < 0 \\ e^{-2x} & \text{if } 0 \leq x \leq \infty \end{cases}$$

Using 10000 simulated values from this distribution, estimate its mean and variance. As in Ross, page 63-66 we obtain the cumulative

distribution function

$$F(x) = \begin{cases} \frac{e^{2x}}{2} & \text{if } x < 0 \\ \frac{2-e^{-2x}}{2} & \text{if } x > 0 \end{cases}$$

The inverse c.d.f gives the generator

$$X = F^{-1}(U) = \begin{cases} \frac{\ln(2U)}{2} & \text{if } U < 1/2 \\ -\frac{\ln(2-2U)}{2} & \text{if } U > 1/2 \end{cases}$$

In Matlab: `U=rand(1,10000); x=-.5*log(2-2*U); v=(U<.5);`
`x(v)=.5*log(2*U(v));` giving `mean(x)` approximately 0 and `var(x)` approximately 0.5 (their true values).

41. 5.6 Let X be an exponential(1) random variable. Give an efficient algorithm for simulating a random variable whose distribution is the distribution of X given that $X < 0.05$. Using 10000 simulated values from this distribution, estimate the mean and variance of this new distribution. A very inefficient algorithm is to generate random variable from the $\text{Exp}(1)$ distribution and then reject all those for which $X > 0.05$. Much more efficient is the inverse transform method, $F(x) = K(1 - e^{-x})$, $0 < x < 0.05$ with $K = 1/(1 - e^{-0.05})$. Then the inverse transform provides $F^{-1}(U) = -\ln(1 - \frac{U}{K}) = -\ln(1 - (1 - e^{-0.05})U)$.
42. 5.7 (The composition method) Suppose it is easy to generate a random variable X_i having a given c.d.f. $F_i(x)$. How could we generate a random variable with c.d.f.

$$F(x) = \sum_{i=1}^n p_i F_i(x)$$

where the values p_i are nonnegative and sum to one? Give an algorithm to generate a random variable X with c.d.f.

$$F(x) = \frac{x + x^3 + x^5}{3}, 0 < x < 1$$

and by simulating 10000 values of this random variable estimate its mean and variance.

This is a mixture and we use the composition method. First generate a random variable I with probability function

$$P(I = i) = p_i$$

Then output the value $X = F_I^{-1}(U)$. For this example the integer I is generated uniformly on the set $\{1, 2, 3\}$ and so, for example `I=3*ceil(rand(1,10000));`
`X=(rand(1,10000)).^(1./I); mean(X)=0.75 var(X)=.0373`

43. 5.9 Give a method to generate a random variable with cumulative distribution function

$$F(x) = \int_0^\infty x^y e^{-y} dy$$

by using composition. Note that if X has conditional c.d.f. given Y $P[X \leq x|Y = y] = x^y$ and Y has a certain distribution, then $F(x)$ is the unconditional c.d.f. of X . Generate $Y = -\ln(U_1)$ and $X = U_2^{1/Y}$ with U_1, U_2 independent $U[0, 1]$.

44. 5.11 Suppose it is easy to generate random variables X_i from each of the c.d.f.s $F_i(x)$. Indicate how to generate a random variable X having as c.d.f.

(a) $\prod_{i=1}^n F_i(x)$

(b) $1 - \prod_{i=1}^n (1 - F_i(x))$.

(Hint: consider the distribution of $\max\{X_i; i = 1, 2, \dots, n\}$ and $\min\{X_i; i = 1, \dots, n\}$ where the X_i are independent with c.d.f. $F_i(x)$. Suppose we generate independent random variables X_i with c.d.f. $F_i(x)$ for each $i = 1, 2, \dots, n$. Then check that the random variable $\max(X_1, \dots, X_n)$ has c.d.f. $F_1(x)F_2(x)\dots F_n(x)$. Therefore we generate X as

$$X = \max(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))$$

It is also easy to show that the minimum

$$X = \min(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))$$

has the c.d.f. in part (b).

45. 5.15 Give an algorithm for generating a random variable from the probability density function

$$f(x) = 30(x^2 - 2x^3 + x^4), 0 \leq x \leq 1$$

and discuss the efficiency of your algorithm. Use acceptance rejection. Note that $30x^2(1-x)^2 \leq \frac{30}{16}$ for all $0 < x < 1$ and so we may use c times a $U[0,1]$ probability density function to dominate $f(x)$. We expect to generate on average $\frac{30}{16}$ or nearly two points (i.e. 4 uniform) in order to produce one random variable with this probability density function.

46. Write a Matlab function which uses the Box-Muller algorithm to generate n pairs (X_1, X_2) of independent normal(0,1) random variables. By simulating 10000 random variables from this algorithm, estimate the value of $E[(X_1 - X_2)^4]$.

The statement

```
u=rand(10000,2); r=sqrt(-2*log(u(:,1))); theta=2*pi*u(:,2);
x=r*[cos(theta) sin(theta)];
```

produces a 10000 by 2 matrix of independent $N(0,1)$ and to estimate the expectation $E[(X_1 - X_2)^4]$ we may use `mean((x(:,1)-x(:,2)).^4)`.

The true value is 12 so this should give a value near 12.

47. Suppose two stocks are presently both worth exactly \$10 and in one month stock 1 will be worth $10 \exp(X_1)$ and stock 2 will be worth $10 \exp(X_2)$ where (X_1, X_2) are both normal random variables with mean 0, variance $\sigma^2 = 0.05$ and the correlation coefficient is $\rho = -0.2$. You have exactly \$10 to spend but we are able to borrow one stock in order to invest in the other. In other words we have the ability to own $y + 1$ shares of stock 1 and $-y$ shares of stock 2 for any value of y (positive or negative).

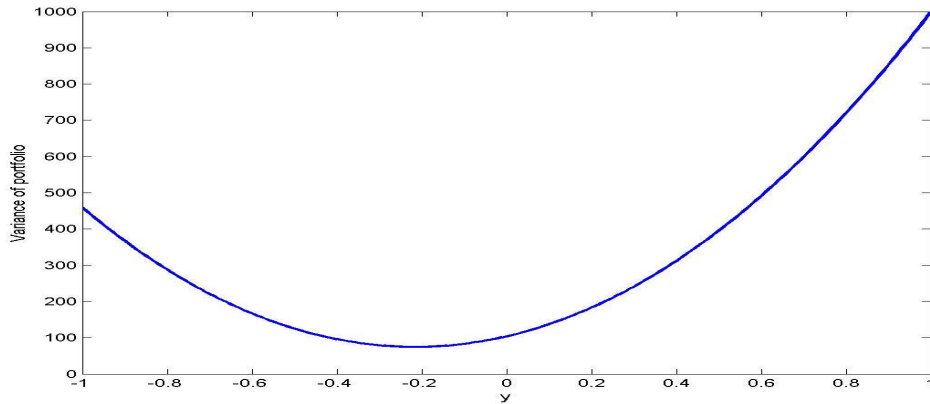


Figure 1:

- Determine whether the expected value of the value of your portfolio $10[(y+1)\exp(X_1) - y\exp(X_2)]$ depends on the value you chose for y .
- For various values of y in the interval $[-1, 1]$ determine the variance of your portfolio $10[(y+1)\exp(X_1) - y\exp(X_2)]$ by simulation.
- What value of y do you recommend if your objective is to achieve the same mean but with the smallest possible variance?

Notice that the expected value of the portfolio after one month is

$$\begin{aligned} 10E[(y+1)\exp(X_1) - y\exp(X_2)] &= 10Ee^{X_1} + y(E[\exp(X_1)] - E[\exp(X_2)]) \\ &= 10E\exp(X_1) \end{aligned}$$

since $\exp(X_1)$ and $\exp(X_2)$ have the same distribution. Therefore the mean does not depend on the value we choose for y . We may generate correlated random variables X_1, X_2 as follows: first generate correlated normal(0,1) random variables and modify to have the correct mean and standard deviation:

```
x=randn(10000,2); rho=-0.2; x(:,2)=rho*x(:,1)+sqrt(1-rho^2)*x(:,2); x=sqrt(0.5)*x;
```

Now consider the variance of the value of possible portfolios for a given values of y , ranging from -5 to 5 .

```
va=[]; for y=-1:.01:1; v=10*((y+1)*exp(x(:,1))-y*exp(x(:,2))); va=[va var(v)]; end
```

we can plot this using

```
plot((-1:.01:1),va); xlabel('y'); ylabel('Variance of portfolio')
```

and this gives a figure like the following: which would appear to indicate that the optimal value of y is around $y = -0.2$. Therefore we should buy 0.8 shares of stock 1 and 0.2 shares of stock 2.

48. 4.10 The negative binomial probability mass function with parameters (r, p) where r is a positive integer and $0 < p < 1$ is given by

$$P[X = j] = \binom{j-1}{r-1} p^r (1-p)^{j-r}, j = r, r+1, \dots$$

- (a) Use the relationship between the negative Binomial and the geometric distribution to simulate from this distribution
(b) Verify the relationship

$$P[X = j+1] = \frac{j(1-p)}{j+1-r} P[X = j]$$

- (c) Use the relationship in (b) to give a second algorithm for generating from the negative binomial distribution
(d) A Negative binomial random variable is generated as the number of trials it takes to generate a total of r successes where trials are independent and the probability of success on each trial is p . Use this fact to generate a negative binomial random variable.

(a) The sum of r independent geometric random variables has a Negative binomial distribution. Therefore since we can generate a Geometric random variable using

$$1 + \left\lceil \frac{\ln(U_i)}{\ln(1-p)} \right\rceil$$

the generator takes the form

$$X = r + \sum_{i=1}^r \left\lceil \frac{\ln(U_i)}{\ln(1-p)} \right\rceil.$$

- (b) this should be a routine calculation

$$\frac{P[X = j+1]}{P[X = j]} = \frac{\binom{j}{r-1} p^r (1-p)^{j+1-r}}{\binom{j-1}{r-1} p^r (1-p)^{j-r}} = \frac{j(1-p)}{j+1-r}$$

- (c) Use inverse transform beginning with $F(r) = P[X = r] = p^r$ and then recursively updating the value of F using $F(j+1) = F(j) + \frac{j(1-p)}{j+1-r} P[X = j]$. (d) Repeatedly generate trials until exactly r successes have been observed. Then count the number of trials required.

49. 4.12 Give two methods for generating a random variable X with

$$P[X = i] = \frac{e^{-\lambda} \lambda^i / i!}{\sum_{j=1}^k e^{-\lambda} \lambda^j / j!}, i = 0, 1, \dots, k$$

one which uses the inverse transform method and the other based on the fact that this distribution is of the form $P[Y = i | Y \in A]$ for some random variable

Y and some set A . Discuss which method you expect to be more efficient for large/small values of k .

The crude method is to repeatedly generate Poisson random variables until we observe a value such that $X \leq k$ and then accept this value. Alternatively we can use inverse transform since if the Poisson cumulative distribution function is $F(x)$ then the cumulative distribution function in this problem takes the form

$$\frac{F(x)}{F(k)}, x = 0, 1, \dots, k.$$

We simply generate a uniform[0,1] random variable U and then search for an integer X which satisfies the inequality

$$\frac{F(X-1)}{F(k)} < U \leq \frac{F(X)}{F(k)}.$$

50. 4.13 Let X be a $\text{Bin}(n, p)$ random variable. Suppose we want to generate a random variable Y whose probability distribution is the same as that of X given that $X \geq k$ for some $k \leq n$. Suppose $\alpha = P[X \geq k]$ has been computed.

- (a) Give the inverse transform method for generating Y
- (b) Give a second method for generating Y
- (c) For what values of α , large or small would the second method be inefficient?

Let X be a binomial random variable with parameters n and p . Suppose that we want to generate a random variable Y whose probability mass function is the same as the conditional mass function of X given that $X \geq k$, for some $k \leq n$. Let $\alpha = P[X \geq k]$ and suppose that the value of α has been computed. (a) Give the inverse transform method for generating Y .

$$P[Y = i] = P[X = i | X \geq k] = \frac{P[X = i, X \geq k]}{P[X \geq k]} = \begin{cases} \frac{P[X=i]}{\alpha} & \text{if } n \geq i \geq k \\ 0 & \text{otherwise.} \end{cases}$$

And this is:

$$P[Y = i] = \begin{cases} \frac{n!}{\alpha i!(n-i)!} p^i (1-p)^{n-i} & \text{if } n \geq i \geq k \\ 0 & \text{otherwise.} \end{cases}$$

We can also use the recursive identity as on page.52 (second edition)

$$P[Y = i+1] = \frac{n-i}{i+1} \frac{p}{1-p} P[Y = i]$$

With i denoting the value currently under consideration, $pr = P[Y = i]$ the probability that Y is equal to i , and $F = F(i)$ the probability that Y is less than or equal to i , the algorithm can be expressed as:

Algorithm

Step 1: Generate a random number U .

Step 2: $c = p/(1-p)$, $i = k$, $pr = \frac{n!}{\alpha k!(n-k)!} p^k (1-p)^{n-k}$, $F = pr$

Step 3: If $(U < F)$, set $Y = i$ and stop.

Step 4: $pr = c(n-i)pr/(i+1)$, $F = F + pr$, $i = i + 1$

Step 5: Go to step 3.

- (b) Give a second method for generating Y we may use Acceptance-Rejection, dominating these probabilities using a binomial distribution. Let Z be a r.v. having a Binomial(n, p) distribution, with probability mass q_j , then let c be a constant such that:

$$\frac{p_j}{q_j} \leq c \quad \text{for all } j \text{ such that } p_j > 0$$

Let $c = 1/\alpha$.

Algorithm

Step 1: Simulate the value of Z .

Step 2: Generate a random number U

Step 3: If $(U < \alpha p_Z/q_Z)$, set $Y = Z$ and stop. Otherwise return to Step 1.

- (c) For small values of α the rejection method in (b) will be inefficient.

51. 4.14 Give a method for generating X from the probability distribution

$$P[X = j] = \begin{cases} 0.11 & j \text{ odd and } 5 \leq j \leq 13 \\ 0.09 & j \text{ even and } 6 \leq j \leq 14 \\ 0 & \text{otherwise.} \end{cases}$$

Using the composition approach:

$$p_j = .55p_j^1 + .45p_j^2$$

where,

$$p_j^1 = .2, \quad j = \{5, 7, 9, 11, 13\} \quad \text{and} \quad p_j^2 = .2, \quad j = \{6, 8, 10, 12, 14\}$$

The probability function is a mixture, $f_1(x)$ is the uniform distribution on the set $\{5, 7, 9, 11, 13\}$ and $f_2(x)$ is the uniform distribution on the set $\{6, 8, 10, 12, 14\}$.

52. 4.15 Suppose that a random variable takes on values

i	1	2	3	4	5	6	7	8	9	10
$P[X = i]$.06	.06	.06	.06	.06	.15	.13	.14	.15	.13

Use the composition method to provide an algorithm for generating the random variable X .

The probability function is a mixture

$$f_4(x) = .3f_1(x) + 0.3f_2(x) + 0.26f_3(x) + .14f_4(x)$$

where $f_1(x)$ is the uniform distribution on the set $\{1, 2, 3, 4, 5\}$, $f_2(x)$ uniform on the two points $\{6, 9\}$, $f_3(x)$ uniform on the points $\{7, 10\}$ and $f_4(x) = 1$ for $x = 8$ (all of the mass is at this one point).

53. 4.16 Give a method for generating values of the random variable X where

$$P[X = i] = \left(\frac{1}{2}\right)^{i+1} + \frac{2^{i-2}}{3^i}, i = 1, 2, \dots$$

This is the mixture of two geometric distributions. With probability $1/2$ generate from $\text{Geom}(1/2)$ and otherwise from $\text{Geom}(1/3)$. If $U_1 < 1/2$ output $X = 1 + \lceil \frac{\ln(U_2)}{1/2} \rceil$ and otherwise if $U_1 > 1/2$ output $X = 1 + \lceil \frac{\ln(U_2)}{2/3} \rceil$.

54. 4.17 Let X be a random variable taking values on the set $\{1, 2, 3, \dots\}$ such that for certain values $\lambda_1, \lambda_2, \dots$

$$P[X = n | X > n - 1] = \lambda_n, n = 1, 2, \dots$$

Show that

$$\begin{aligned} P[X = n] &= \lambda_n(1 - \lambda_{n-1}) \dots (1 - \lambda_1) \text{ for } n \geq 2 \\ P[X = 1] &= \lambda_1 \end{aligned}$$

and that

$$\lambda_n = \frac{p_n}{1 - \sum_{i=1}^{n-1} p_i}, n = 2, 3, \dots$$

Show that $p_1 = \lambda_1$ and $p_n = (1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{n-1})\lambda_n$.

We can verify the formula by induction.

I) We verify for $n = 1$

$$\lambda_1 = \frac{p_1}{1 - 0} = p_1$$

II) We assume for n :

$$p_n = (1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{n-1})\lambda_n$$

III) We prove for $n = n + 1$

$$\begin{aligned} \lambda_{n+1} &= \frac{p_{n+1}}{1 - \sum_{j=1}^n p_j} \\ &= \frac{p_{n+1}}{1 - \sum_{j=1}^{n-1} p_j - p_n} \\ &= \frac{p_{n+1}}{(1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{n-1}) - p_n} \\ &= \frac{p_{n+1}}{(1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{n-1}) - (1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{n-1})\lambda_n} \\ &= \frac{p_{n+1}}{(1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_{n-1})(1 - \lambda_n)} \end{aligned}$$

So,

$$p_{n+1} = (1 - \lambda_1)(1 - \lambda_2) \dots (1 - \lambda_n)\lambda_{n+1}$$

Show that the value of X when the above stops has the desired mass function.

Suppose that the algorithm stops at $X = m$ then, $U_m < \lambda_m$ and $U_{m-1} \geq \lambda_{m-1}$, i.e. $U_{m-1} < 1 - \lambda_{m-1}$, $U_{m-2} < 1 - \lambda_{m-2}$ ■. So we have:

$$P(X = m) = P(U_m < \lambda_m)P(U_{m-1} \geq \lambda_{m-1}) \dots P(U_1 \geq \lambda_1) = p_m$$

Suppose that X is a geometric random variable with parameters p . Determine the values λ_n , $n \geq 1$. Explain what the above algorithm is doing in this case and why its validity is clear.

If X is a geometric r.v. then

$$p_n = p(1-p)^{n-1} \quad n \geq 1$$

Then we have:

$$\begin{aligned} \lambda_n &= \frac{p_n}{1 - \sum_{j=1}^{n-1} p_j} \\ &= \frac{p(1-p)^{n-1}}{1 - \sum_{j=1}^{n-1} p_j} \\ &= \frac{p(1-p)^{n-1}}{1 - p[1 + (1-p) + (1-p)^2 + \dots + (1-p)^{n-2}]} \\ &= \frac{p(1-p)^{n-1}}{1 - p \frac{1-(1-p)^{n-1}}{1-(1-p)}} \\ &= p \end{aligned}$$

55. 5.16 Consider generating a $N(0, 1)$ random variable Z by first generating the absolute value $|Z|$, having probability density function

$$f(x) = \sqrt{\frac{2}{\pi}} \exp(-x^2/2), x > 0$$

and then assigning it a random sign. In order to generate a random variable from the probability density function f , we use acceptance rejection, dominating f using an exponential density $g(x) = \lambda e^{-\lambda x}$, $x > 0$. Show that the choice of λ which provides the most efficient acceptance-rejection method corresponds to $\lambda = 1$. Determine how many iterations are required on average to generate a single value of $|Z|$.
Let

$$c = c(\lambda) = \max_x \frac{f(x)}{g_\lambda(x)}$$

where

$$f(x) = \sqrt{\frac{2}{\pi}} e^{-x^2/2} \text{ and } g_\lambda(x) = \lambda e^{-\lambda x}, \text{ for } x > 0$$

Then it is easy to show by calculus that

$$c(\lambda) = \frac{1}{\lambda} \sqrt{\frac{2}{\pi}} e^{\lambda^2/2}$$

The best choice of λ is the choice which minimizes the value of $c(\lambda)$. Again using calculus we can show that this corresponds to $\lambda = 1$.

56. 5.21 Consider a job that consists of a number of a total of k stages that must be done in sequence. The time required to complete stage i is an exponential random variable with parameter λ_i and then after completing stage i , the worker goes on to stage $i + 1$ with probability α_i , otherwise the worker stops working with probability $(1 - \alpha_i)$. Let X be the total amount of time the worker spends on the job. Write an algorithm for generating the value of X . Note that if $P[Y = j] = \alpha_1 \alpha_2 \dots \alpha_{j-1} (1 - \alpha_j)$ then $P[Y > j] = \alpha_1 \alpha_2 \dots \alpha_j$. For independent $U[0,1]$ random variables U_i , define a random variable $Y = j$ if and only if $U_i < \alpha_i$ for all $i < j$ and $U_j \geq \alpha_j$. Thus Y is the smallest index such that $U_Y \geq \alpha_Y$. Alternatively we can use inverse transform to generate Y with $F(y) = 1 - \prod_{i=1}^y \alpha_i$ (this is in fact the same method). Then Y is the number of stages required. The amount of time required is the sum of Y exponentially distributed random variables

$$X = - \sum_{i=1}^Y \frac{\ln(V_i)}{\lambda_i}, \text{ where } V_i \text{ are independent } U[0, 1].$$

57. 5.23 Buses arrive at a sporting event according to a Poisson process with rate 5 per hour. Each bus is equally likely to carry 20, 21, ..., 40 fans with the numbers in different buses being independent. Write an algorithm to simulate the arrival of fans in the time interval $0 < t < 1$. Notice that

$$\lambda(t) \leq 7 \text{ for all } t > 0.$$

Generating a sequence of independent bus arrival times T_i independent exponential(5) and then at each generate the number of passengers. Record at event times in the vector ET the total number of passengers NA that have arrived.

- (a) $t = 0; NA = []; ET = [];$
- (b) while ($t < 1$)
- (c) $t = t - \frac{\ln(\text{rand})}{5};$
- (d) $ET = [ET \ t]; NA = [NA \ 19 + \text{ceil}(21 * \text{rand})];$
- (e) end

58. 5.24 Give an efficient algorithm for generating points on the interval $0 < t < 10$ of a Poisson process which has intensity function

$$\lambda(t) = \begin{cases} \frac{t}{5} & \text{for } 0 < t < 5 \\ 1 + 5(t - 5) & \text{for } 5 < t < 10 \end{cases}$$

See Ross pages 78-79 for a method to generate the Poisson process. Generate a Poisson process on the interval $0 < t < 5$ with intensity 1. Thin using the above thinning method to obtain the process on $0 < t < 5$ which has intensity function $\lambda(t) = \frac{t}{5}$. For $5 < t \leq 6$, again generate the Poisson process with intensity $\lambda(t) = 1 + 5(t - 5)$ by thinning a process with constant intensity 6. Similarly for $6 < t \leq 7$, thin a process with constant intensity 11, etc. This is more efficient than using a constant intensity of 26 and thinning to obtain the process over the whole interval $5 < t < 10$.

59. Give two algorithms for generating observations from a distribution with probability density function

$$f(x) = \frac{(x-1)^3}{20}, \text{ for } 2 \leq x \leq 4.$$

Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. Which algorithm is faster?

The simplest method is to use inverse transform, since in this case the cumulative distribution function

$$F(x) = \frac{(x-1)^4}{80}, \text{ for } 2 \leq x \leq 4$$

is easily inverted. An alternative might be to use acceptance-rejection dominating the density with a function like $c(x-1)$ for $2 \leq x \leq 4$ but acceptance-rejection is bound to be slower.

60. Give two different algorithms for generating observations from a distribution with a density function of the form $f(x) = cx^3e^{-x/2}$ for $x > 0$ and appropriate constant c . Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. Which algorithm is faster? This is a gamma(4,2) distribution and the constant is

$$c = \frac{1}{\Gamma(4)2^4} = \frac{1}{96}$$

so we can either use the sum of four independent exponential random variables (each Gamma(1,2))

$$X = -2 \sum_{i=1}^4 \ln(U_i)$$

or we can use acceptance-rejection. Try dominating this probability density function using an exponential random variable with the same mean, so $g(x) = \frac{1}{8} \exp(-x/8), x > 0$. Note that

$$\frac{f(x)}{g(x)} = \frac{8x^3e^{-x/2}}{96 \exp(-x/8)} = \frac{1}{12}x^3e^{-\frac{3}{8}x}$$

whose maximum value $\frac{128}{3}e^{-3} = 2.1242$ occurs for $x = 8$. Therefore the value of c is 2.1242. The algorithm is

Step 1. Generate $X = -8 \ln(U_1)$.

Accept X if $\frac{1}{12}X^3 e^{-\frac{3}{8}X} \leq cU_2$, and otherwise return to step 1.

There are other functions one can use to dominate more closely the gamma density function (see the notes for example) We might use a function of the form

$$g(x) = \frac{pkx^{p-1}}{(k+x^p)^2}$$

and try and find a suitable value for p, k . One suggestion is to choose these values so that the mode of the functions f and g match, giving $k = \alpha^p$ and $p = \sqrt{2\alpha - 1}$. In this case

$$f(x) \leq cg(x)$$

with $c = \frac{2\alpha^\alpha}{p\Gamma(\alpha)e^\alpha}$. This is close to one for $\alpha = 4$ and this method is very efficient.

61. Give a precise algorithm for generating observations from a discrete distribution with $P[X = j] = (2/3)(1/3)^j$; $j = 0, 1, \dots$. Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. Compute as well the sample variance. How large would the simulation need to be if we wanted to estimate the mean within 0.01 with a 95% confidence interval?

This is a variation of the geometric distribution. In this case, X counts the number of failures we generated before obtaining a single success if the probability of a success on each trial is $2/3$. In other words, $X = Y - 1$ where Y has a geometric distribution with $p = 2/3$. Therefore we can generate, by inverse transform,

$$X = \left\lfloor \frac{\ln(1-U)}{\ln(1-p)} \right\rfloor$$

where $\lfloor \cdot \rfloor$ represents the integer part.

62. Give a precise algorithm for generating observations from a distribution which has probability density function $f(x) = x^3$, $0 < x < \sqrt[4]{2}$. Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. Determine the standard error of the sample mean. How large would the simulation need to be if we wanted to estimate the mean within 0.01 with a 95% confidence interval? This is a simple inverse transform problem. Note that

$$F(x) = \frac{x^4}{4}, \quad 0 < x < \sqrt[4]{2}$$

and so the inverse transform is

$$X = (4U)^{1/4}.$$

The mean is $E(X) = \int_0^{\sqrt[4]{2}} xf(x)dx = \int_0^{\sqrt[4]{2}} x^4 dx = \frac{4}{5}\sqrt[4]{2} = 1.1314$. and the standard deviation of X is around 0.23. In order to estimate the mean within

0.01 with a 95% confidence interval we need to do n simulations where n satisfies

$$\begin{aligned} 2 \times \frac{0.23}{\sqrt{n}} &\leq 0.01 \\ n &\geq 46^2 = 2116 \end{aligned}$$

63. Consider independent random variables X_i with c.d.f.

$$\begin{aligned} F_i(x) &= x^2, & i = 1 \\ &= \frac{e^x - 1}{e - 1}, & i = 2 \\ &= xe^{x-1}, & i = 3 \end{aligned}$$

for $0 < x < 1$. Explain how to obtain random variables with cumulative distribution functions

$$G(x) = \prod_{i=1}^3 F_i(x) \quad \text{and} \quad G(x) = 1 - \prod_{i=1}^3 (1 - F_i(x)).$$

We need to generate three random variables X_i having c.d.f. $F_i(x)$ for $i = 1, 2, 3$ and then use $\max(X_1, X_2, X_3)$ as a random variable having c.d.f. $\prod_{i=1}^3 F_i(x)$ and $\min(X_1, X_2, X_3)$ having c.d.f. $1 - \prod_{i=1}^3 (1 - F_i(x))$. The random variables X_1, X_2 can be generated by inverse transform. The random variable X_3 can be generated using acceptance-rejection. We can use the density function $g(x) = \frac{2}{3}(x+1)$ to dominate x since the corresponding density satisfies

$$f_3(x) = (x+1)e^{x-1} \leq (x+1) \quad \text{on } 0 < x < 1$$

Since in this case $c = 3/2$ we can expect reasonable efficiency (only an average of 1 point in 3 is rejected).

64. Evaluate the following integral by simulation. Give two different methods

$$\int_0^1 (1-x^2)^{3/4} dx.$$

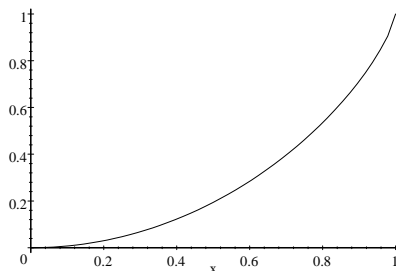
The true value is $\frac{1}{2} \text{Beta}\left(\frac{1}{2}, \frac{7}{4}\right) = 0.71888$. We may either use crude Monte Carlo, i.e. evaluate

$$\frac{1}{n} \sum_{i=1}^n (1 - U_i^2)^{3/4}$$

or use antithetic random numbers

$$\frac{1}{2n} \sum_{i=1}^n [(1 - U_i^2)^{3/4} + (1 - (1 - U_i)^2)^{3/4}].$$

The graph of the function $f(x) = 1 - (1 - x^2)^{3/4}$ shows that the function is monotone on the interval $[0,1]$ and so the method of antithetic random numbers should provide some variance reduction.



65. Evaluate the following integral by simulation:

$$\int_0^2 x^{3/2}(4-x)^{1/2} dx.$$

This equals

$$E(2X^{3/2}(4-X)^{1/2})$$

where $X = 2U$ has the uniform[0,2] distribution. Therefore the simulated value is

$$\frac{2}{n} \sum_{i=1}^n (2U_i^{3/2})(4-2U_i)^{1/2}$$

The true value is $-\frac{8}{3} + 2\pi = 3.6165$

66. Find by simulation the area of the region $\{(x, y); -1 < x < 1, y > 0, \sqrt{1-2x^2} < y < \sqrt{1-2x^4}\}$. The boundaries of the region are graphed in Figure 1. Which method do you expect to be more efficient, generating random points in the rectangle $(-1 < x < 1, 0 < y < 1)$ and counting the proportion that lie in the region between these curves or generating a random value of x in the interval $[-1, 1]$ and using a multiple of the average of the values of $\sqrt{1-2x^4} - \sqrt{1-2x^2}$? We begin by generating values in the rectangle and counting what proportion of these points fall in the region between the two graphs. Define $(X_i, Y_i) = (2U_i - 1, V_i)$ where both U_i, V_i are independent $U[0, 1]$. Then the crude Monte Carlo estimator is the average of the indicator random variables

$$\hat{\theta}_{crude} = \frac{1}{n} \sum_{i=1}^n I(\sqrt{1-2X_i^2} < Y_i < \sqrt{1-2X_i^4}) = \frac{1}{n} \sum_{i=1}^n I_i$$

Notice however that if you were told the value of X_i , then the conditional expectation

$$E[I_i|X_i] = \sqrt{1-2X_i^4} - \sqrt{1-2X_i^2}$$

and since the conditional expectation $E(I_i|X_i)$ has the same mean as does I_i but smaller variance, a second estimator with the same expected value as $\hat{\theta}_{crude}$ but

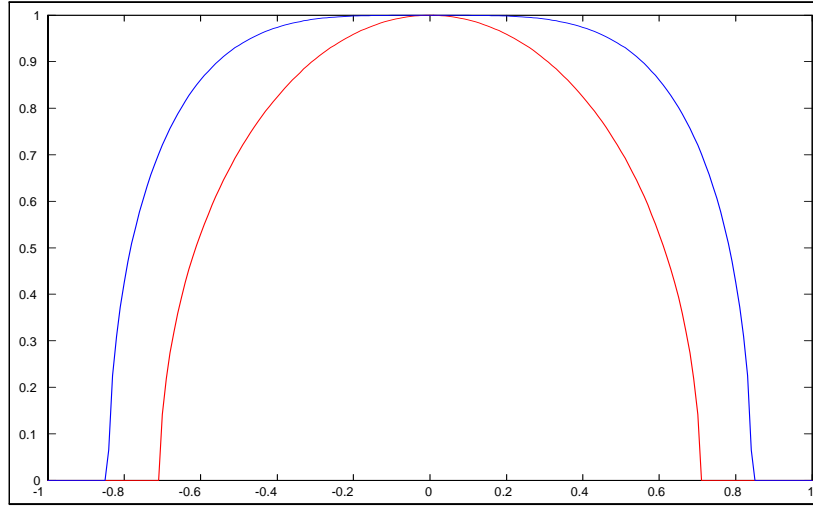


Figure 2:

with smaller variance is

$$\frac{1}{n} \sum_{i=1}^n (\sqrt{1 - 2X_i^4} - \sqrt{1 - 2X_i^2}).$$

67. Briefly indicate an efficient algorithm for generating one random variable from each of the following distributions and generate such a random variable using one or more of the uniform[0,1] random numbers.

U_i :	0.794	0.603	0.412	0.874	0.268	0.990	0.059	0.112	0.395
---------	-------	-------	-------	-------	-------	-------	-------	-------	-------

- (a) $X \sim U[-1, 2]$. $3U - 1$
- (b) a random variable X with probability density function $f(x) = \frac{3}{16}x^{1/2}$, $0 < x < 4$ $X = 4U^{2/3}$
- (c) A discrete random number X having probability function $P[X = x] = (1 - p)^x p$, $x = 0, 1, 2, \dots, p = 0.3$. Inverse transform $X = \left\lceil \frac{\ln(U)}{\ln(1-p)} \right\rceil$
68. Briefly indicate an efficient algorithm for generating one random variable from each of the following distributions and generate such a random variable using one or more of the uniform[0,1] random numbers.

U_i :	0.794	0.603	0.412	0.874	0.268	0.990	0.059	0.112	0.395
---------	-------	-------	-------	-------	-------	-------	-------	-------	-------

- (a) A random variable X with the normal distribution, mean 1 and variance 4. Generate (Z_1, Z_2) using Box Muller and use $(X, Y) = 1 + 2(Z_1, Z_2)$

(b) A random variable X with probability density function

$$f(x) = cx^2e^{-x}, \quad 0 \leq x < 1$$

for constant $c = 1/(2 - 5e^{-1})$. You may either use acceptance rejection, generating from density such that $g(x) = 3x^2, 0 < x < 1$ or generate gamma(3,1) variates repeatedly (see the gamma distribution generators) until one is less than 1.

(c) A random variable X with the following probability function:

x	0	1	2	3	
$P[X = x]$	0.1	0.2	0.3	0.4	

If $U < .1$, put $X = 0$, if $.1 < U < .3$, put $X = 1$, etc. (Inverse transform).

69. Write two algorithms for generating Gamma(20,1) random variables. One should use the sum of 20 independent exponential random variables and the other uses acceptance rejection with a dominating function of the form

$$\text{constant} \times \frac{x^{p-1}}{(c + x^p)^2}$$

for sensible values of c and p . Compare the speed of the two algorithms. We wish to generate from a Gamma distribution having mean 20 and variance 20. If I let

$$g(x) = \text{constant} \times \frac{x^{p-1}}{(c + x^p)^2}$$

can we find parameters p and c so that we have approximately the same location and shape of distribution? Note that

$$\int_0^z \frac{x^{p-1}}{(c + x^p)^2} dx = -\frac{1}{p(c + z^p)} + \frac{1}{pc}$$

and so

$$g(x) = g(x) = cp \frac{x^{p-1}}{(c + x^p)^2}$$

and the cumulative distribution function is

$$G(z) = 1 - \frac{c}{(c + z^p)}.$$

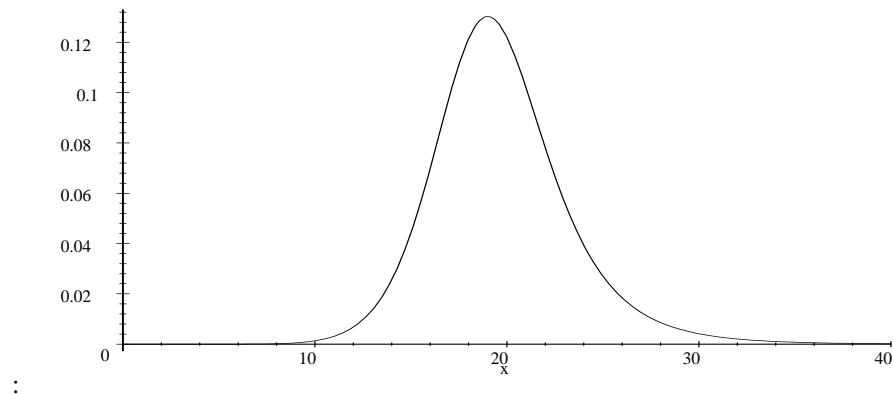
This is easy to invert. We find appropriate values of the parameters c, p , for example, by setting the mode of the distribution (the value of x where the density reaches its maximum) equal to the mode of the Gamma (=19 in this case). To this end we must solve

$$\frac{1}{2} - \frac{1}{p} = \frac{19^p}{c + 19^p}$$

and this can be used to determine c as a function of p .

$$\begin{aligned} c &= 19^p \frac{p+1}{p-1} \\ p &= 10 \end{aligned}$$

We now try several values of p to see if we get approximately the right shape of the distribution. For example in the case $p = 10$ we obtain a probability density function that resembles the Gamma



70. *Haggis Elevators* has a total of four elevators installed in MC. A typical elevator breaks down once every 10 days. One repairman can repair an elevator in an average of 4 days (from the time that the repair is ordered). Currently there is one repairmen on call. Service times and the time between breakdowns are exponentially distributed.

- (a) Draw a state diagram including transition rates for the Markov Chain. The state is the number of elevators operating. Breakdown rate= $1/10$ per operating elevator. Service rate= $1/4$.

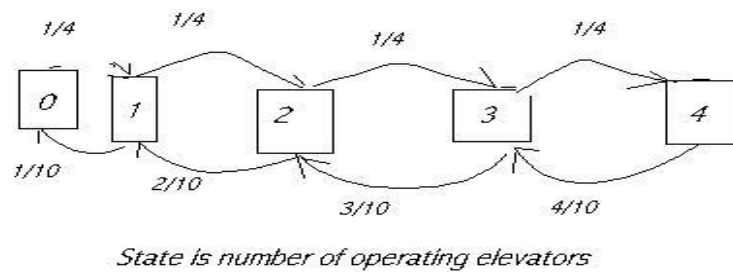


Figure 3:

- (b) Assuming that the system begins with all four elevators operating, use the uniform numbers and their negative logarithms listed below to simulate the operation of the system for the first 4 events.

U_i : 0.019 0.821 0.445 0.180 0.792 0.922 0.738 0.240
 $-\ln(U_i)$: 3.99 0.20 0.81 1.71 0.23 0.08 0.30 1.43

Define the variables below and fill in the following table:

System state=Number of elevators operating $\{0,1,2,3,4\}$

Event List=(t_B, t_S) =(time of next breakdown, time of next service completion).

Clock Time	System State	Event List
0	4	(9.975, ∞)
9.975	3	(12.675, 10.775)
10.775	4	(15.05, ∞)
15.05	3	(15.32, 15.97)
15.32	2	(16.82, 15.97)

- (c) Run a simulation and compare the observed proportion of time each of the 5 states are occupied with the expected time when the system is operating at steady state. (you need to solve for the equilibrium distribution)
71. A communications network transmits binary digits, 0 or 1. There is probability q that a digit transmitted will be incorrectly received at the next stage. If X_0 denotes the binary digit originally transmitted, X_n the digit after it has been retransmitted n times, simulate the process for $n = 1, 2, \dots, 10000$ beginning both with $X_0 = 1$ and $X_0 = 0$ and for various values of q (e.g. $q = 0.1, 0.2, \dots, 0.5$). Record the proportion of time that the process spends in each of its two states 0 and 1. *The proportion of time that the chain is in state 0 should be close to the equilibrium probability of state 0. This is obtained by solving the equations $\pi'P = \pi'$ giving solution $(1/2, 1/2)$ no matter what the value of q (if $0 < q < 1$).*
- (b) Given that 0 is originally transmitted ($X_0 = 0$), what is the probability that the first error in retransmission (i.e. $X_n = 1$ for the first time) occurs on the n 'th retransmission. What is the *expected time* before the chain returns to its original state. *The probability of the first error occurring after n transmissions is given by the geometric distribution $(1 - q)^{n-1}q$. Expected time to first change of state = $1/q$. Expected time till second change of state and it returns to the original one = $2/q$.*
72. Two mechanics are employed to service 3 aircraft. All aircraft not undergoing repair are in use and the time until breakdown of any one aircraft while in use has an exponential distribution with expected value $1/\lambda$. One mechanic is assigned to each aircraft requiring repair, and the time until this repair is completed is exponential with expectation $1/\mu$. Let X_t be the number of aircraft requiring repair at time t (Note that when $X_t = 3$, only two aircraft are being repaired and

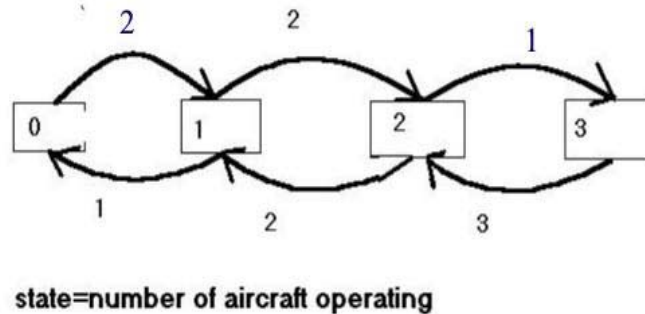


Figure 4:

one is waiting to begin repair). Draw a state diagram including transition rates for the Markov Chain. Determine by simulation the answers to the following questions. Try and indicate some measure of the precision of your estimate and compare where possible with the theoretical equilibrium values.

- (a) Suppose $\mu = \lambda = 1$. If the company loses \$1,000 per unit time that only one aircraft is operating and \$5,000 for each unit of time when no aircraft are operating, what is the long run average cost to the company per unit time?
- (b) Suppose $\mu = \lambda = 1$. What is the average length of time that an aircraft spends waiting before beginning repair?

The state transition diagram with rates of transitions are as in the following diagram: **FIX THE SOLUTION**. Solving for the equilibrium distribution by equating flow in and flow out for the various states, we obtain the equations

$$\begin{aligned} 2\pi_0 &= \pi_1 \\ 3\pi_1 &= \pi_0 + 2\pi_2 \\ 3\pi_2 &= 2\pi_1 + 3\pi_3 \end{aligned}$$

and solving these equations we obtain

$$\pi = \pi_0(1, 1, 5/2, 7/6)$$

so solving for π_0 gives $\pi_0 = \pi_1 = \pi_2 = 3/11$ and $\pi_3 = 2/11$. Since the proportion of time that there are no aircraft operating is π_0 and the proportion of time there is only one operating π_1 , the cost per unit time is $\$1000\pi_1 + \$5,000\pi_0$. The average length of time before an aircraft begins repair depends on the number of aircraft in the

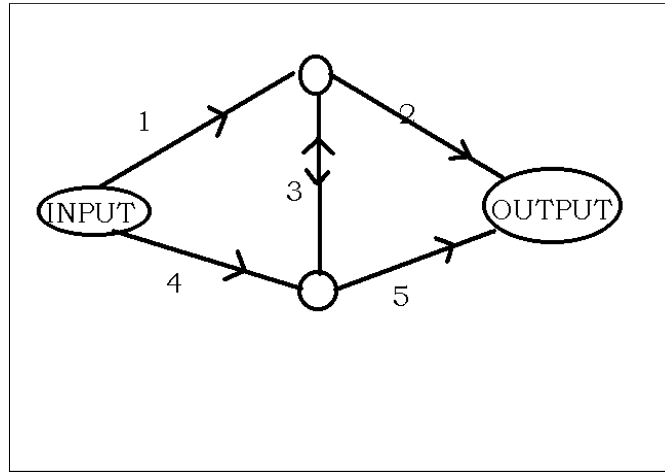


Figure 5:

73. Consider the following network of five components (components are represented by directed arcs in Figure 3).The network operates as long as there is flow permitted from the input (source) to the output (sink). The arrows indicate the direction of flow possible. Find by simulation the expected lifetime of this system if the individual components all have independent lifetimes

(a) that are exponential with mean 1.

(b) that are uniform on the interval $[0,2]$. Suppose the lifetimes of the components are T_1, \dots, T_5 . Then the lifetime of the system is $\max(Y_1, Y_2, Y_3, Y_4)$ where $Y_1 = \min(T_1, T_2)$, $Y_2 = \min(T_1, T_3, T_5)$, $Y_3 = \min(T_4, T_5)$, $Y_4 = \min(T_4, T_3, T_5)$.

$L=[]$;

for $i=1:10000$

$T=2*\text{rand}(1,5)$; (for part b) or $T = -\log(\text{rand}(1,5))$ (for part (a)

$L=[L \max(\min(T([1 \ 2])), \min(T([1 \ 3 \ 5])), \min(T([4 \ 5])), \min(T([4 \ 3 \ 2])))]$;

end

$\text{mean}(L)$; $\text{sqrt}(\text{var}(L)/\text{length}(L))$; $\text{hist}(L,50)$

74. The interarrival times between consecutive buses at a certain bus stop are independent $\text{Beta}(1/2, 1/2)$ random variables with probability density function

$$f(x) = \frac{1}{\pi \sqrt{x(1-x)}}, \quad 0 < x < 1.$$

starting at clock time $t = 0$. You arrive at the bus stop at time $t = 10$. Determine by simulation the expected time that you will have to wait for the next bus. Is

it more than $1/2$? Repeat when the interarrival times have the distribution of Y^2 where Y is an exponential with parameter $\lambda = 10$. Explain how this is possible.

$F(y) = \int_0^y \frac{1}{\pi} \frac{1}{x(1-x)} dx = \frac{2 \arcsin(2y-1) + \pi}{2\pi}$. Use inverse transform to generate interarrival times. $F^{-1}(U) = \frac{1}{2}(1 + \sin \pi(U - \frac{1}{2}))$.

```
function q74
% simulated arrival at a bus stop-question 74
%X=.5*(1+sin(pi*(rand(1,10000)-.5))); %beta distribution
%X=-.5*log(rand(1,100000)); % exponential distribution
X=(-.1*log(rand(1,100000))).^2; %wweibull distribution
Z=[];
while (sum(X)>10)
Y=cumsum(X);
i=1+sum(Y<10); Z=[Z Y(i)-10];
X=X((i+1):length(X));
end
disp(['your average wait/average time between buses=' num2str(mean(Z)/mean(X))])
```

This gives a mean of around .3711. The average wait is less than the average time between buses. However, if we use a squared exponential variate in place of X , i.e. $X = E^2$ where E is exponential with mean $(1/10)$, and replace the generation of X by $X=(-.1\log(\text{rand}(1,n))).^2$; we discover that your average wait is 3-4 times the average time between buses. When you arrive at a random time, the longer interarrival times are more likely to be those in which you arrive. This is called length-biased sampling—longer intervals are more likely selected than shorter ones.*

75. Ross 7.8 If we add random numbers until their sum exceeds 1, then the expected number added is e . In other words if

$$N = \min(n; \sum_{i=1}^n U_i > 1)$$

then $E(N) = e$. Using a simulation estimate the variance of N and determine the number of observations of N required in order to estimate e accurately to within 0.001 (the width of a 95% confidence interval is around 0.002). This is the same simulation conducted earlier using the function

```
function N=generateN(m)
N=[];
for i=1:m
V=cumsum(rand(1,100));
N=[N 1+sum(V<1)];
```

end

Running this function as with `N=generateN(1000); mean(N); var(N);` produces an estimator of the mean around 2.71 and the variance around .7627. Since the 95% confidence interval takes the approximate form

$$\bar{X} \pm 2S/\sqrt{n}$$

we have

$$2S/\sqrt{n} \leq .001$$

or

$$n \geq \left(\frac{2S}{.001}\right)^2 = \frac{4 \times .7627}{(.001)^2}$$

or $n \geq 3,050,800$.

76. Ross 7.9 Consider a sequence of random $U[0, 1]$ numbers and let M be the first one that is less than its predecessor:

$$M = \min\{n; U_1 \leq U_2 \leq \dots \leq U_{n-1} > U_n\}$$

- (a) Argue that $P[M > n] = \frac{1}{n!}, n = 0, 1, \dots$
- (b) Use the identity $E(M) = \sum_{n=0}^{\infty} P[M > n]$ to show $E(M) = e$.
- (c) Use 1000 simulation runs and part b to estimate e with a 95% confidence interval.

77. Ross 7.14 If $n = 2$ and $X_1 = 1, X_2 = 3$ determine what is the bootstrap estimator of $\text{var}(S^2)$. The bootstrap estimator is obtained by taking random samples of size 2 with replacement from the values $\{1, 3\}$ repeatedly and computing the value of the sample variance S^2 . There are only three possible samples with probabilities as below

Sample	Probability	S^2
$\{1, 1\}$	1/4	0
$\{1, 3\}$ (either order)	1/2	2
$\{3, 3\}$	1/4	0

Therefore the bootstrap estimator of the variance of S^2 will be the variance of a random variable which takes the values 0 and 2 each with probability 1/2. This is therefore 1.

78. Ross 7.15 If $n = 15$ and we observe the data

5, 4, 9, 6, 21, 17, 11, 20, 7, 10, 21, 15, 13, 16, 8

obtain a bootstrap estimator of $\text{var}(S^2)$ using 1000 bootstrap samples. We use the following function

```
function s=sample(x,n)
```

```
%takes a sample of size n with replacement from the vec-
  tor x
s=unidrnd(length(x),1,n);          s=x(s);
  and then obtain the bootstrap estimator as follows
x=[5,4,9,6,21,17,11,20,7,10,21,15,13,16,8];
S2=[];
for i=1:1000
    bootx=sample(x,length(x));      S2=[S2 var(bootx)];
end
disp(['bootstrap estimate of var(S^2) = ' num2str(var(S2))])

and this gives a value around 64.7
```

79. Use both crude and antithetic random numbers to integrate the function

$$\int_0^1 \frac{e^u - 1}{e - 1} du.$$

What is the efficiency gain attributed to the use of antithetic random numbers?

First for a crude monte-carlo estimator:

```
t1=cputime U=rand(1,200000); % crude Monte Carlo
F=(exp(U)-1)/(exp(1)-1); % value of 200000 values of the
function
m=mean(F) % average value
S1=var(F)/length(F) % estimated variance of estimator
t1=[t1 cputime];
```

Now we use 100000 antithetic pairs for the same number of function evaluations. *****

```
U=rand(1,100000);
F=.5*(exp(U)-1)/(exp(1)-1)+.5*(exp(1-U)-1)/(exp(1)-1);
mean(F)
S2=var(F)/length(F) ;
t1=[t1 cputime];
display(['times ' mat2str(diff(t1))]);
display(['efficiency= ' num2str(S1/S2)])
```

This gives means around .4180 and variances $S1 = 4.1 \times 10^{-7}$ and $S2 = 1.32 \times 10^{-8}$ for a ratio (the efficiency gain due to antithetic) of around 31. Moreover the antithetic takes about 30% less time.

80. Under what conditions on f does the use of antithetic random numbers completely remove the variance of the Monte-Carlo estimator? i.e. When is $\text{var}(f(U) + f(1 - U)) = 0$? If the variance of a random variable is 0 then the random

variable must be constant so $f(U) + f(1 - U) = c$ for all $0 < U < 1$. This implies that $f(\frac{1}{2} + x) = c - f(\frac{1}{2} - x)$ or that f is an odd function about the point $(1/2, c)$.

81. Show that if we use antithetic random numbers to generate two normal random variables X_1, X_2 , having mean μ and variance σ^2 , this is equivalent to setting $X_2 = 2\mu - X_1$. In other words, it is not necessary to use the inverse transform method to generate normal random variables in order to permit the use of antithetic random numbers. Use antithetic random numbers to estimate the value of a European call option $E[e^{-rT} \max(10e^X - 10, 0)]$ for $r = .10, T = .5, \sigma = .3$. You need to show that if $F(x)$ is the normal c.d.f. with mean μ and variance σ^2 and if $X_1 = F^{-1}(U)$ then $F^{-1}(1 - U) = 2\mu - X_1$. Notice that if we let Φ be the standard normal c.d.f. then $F(x) = \Phi(\frac{x-\mu}{\sigma})$ so that $F(2\mu - X_1) = \Phi(\frac{\mu - X_1}{\sigma}) = 1 - \Phi(\frac{X_1 - \mu}{\sigma}) = 1 - F(X_1) = 1 - U$.
82. Show that it is better to use U^2 than U as a control variate when estimating $E(\sqrt{1 - U^2})$. The necessary covariance may be estimated by simulation. We need to compare the variance of $(\sqrt{1 - U^2} - \beta_1 U^2)$ and the variance of $(\sqrt{1 - U^2} - \beta_2 U)$ where these coefficients are chosen optimally

$$\begin{aligned}\beta_1 &= \frac{\text{cov}(\sqrt{1 - U^2}, U^2)}{\text{var}(U^2)} = \frac{-.0656}{.0891} = -.736 \\ \beta_2 &= \frac{\text{cov}(\sqrt{1 - U^2}, U)}{\text{var}(U)} = \frac{-.0595}{.0833} = -.714\end{aligned}$$

Gives $\text{var}(\sqrt{1 - U^2} - \beta_1 U^2) = .0016$, $\text{var}(\sqrt{1 - U^2} - \beta_2 U) = .0075$ whereas crude estimator has variance .05. U^2 is a better control variate and produces an efficiency gain of around 31.

83. Use a stratified random sample to integrate the function

$$\int_0^1 \frac{e^u - 1}{e - 1} du.$$

What do you recommend for intervals (two or three) and sample sizes? What is the efficiency gain over crude Monte Carlo?

By experimentation, two intervals such as $[0 .55], [.55 1]$ results in optimal sample sizes (for a total sample of 100000) around 47400 and 52600 respectively and provides an efficiency gain of around 4 over crude monte carlo. Three intervals e.g. $[0 .35], [.35 .75], [.75 1]$ with sample sizes about 25229, 48686, 26083 provides an efficiency gain of around 8 over crude monte carlo.

84. Use a combination of stratified random sampling and an antithetic random number in the form

$$\frac{1}{2}[f(U/2) + f(1 - U/2)]$$

to integrate the function

$$f(u) = \int_0^1 \frac{e^u - 1}{e - 1} du.$$

What is the efficiency gain?

The true value of the integral is about 0.41802. Using crude, the estimator has variance about 0.0822/100000 for a crude monte carlo sample of 100000. The average of 50,000 estimators of the form $\frac{1}{2}[f(U/2) + f(1-U/2)]$ has variance .0013/50000. The efficiency gain is around 32.

85. In the case $f(x) = \frac{e^x - 1}{e - 1}$, use $g(x) = x$ as a control variate to integrate over $[0,1]$. Show that the variance is reduced by a factor of approximately 60. Is there much additional improvement if we use a more general quadratic function of x ?

The variance of the control variate estimator (simple version) is determined by $\text{var}(f(u) - g(u))$ where $g(u) = u$. The efficiency gain is .0822/.0014 or about 58. What if we use a general quadratic function as control variate. We might crudely choose a quadratic function that takes the same values i.e. 0, .3775, and 1 as $f(x)$ at the points $x = 0, 1/2$, and 1. $g(x) = .51x + .49x^2$. Then $\text{var}(f(u) - g(u))$ is approximately .0000317 and the efficiency gain is .0822/.0000317 or around 2600!

86. In the case $f(x) = \frac{e^x - 1}{e - 1}$, consider using $g(x) = x$ as a control variate to integrate over $[0,1]$. Note that regression of $f(U)$ on $g(U)$ yields $f(U) - E(f(U)) = \beta[g(U) - E(g(U))] + \varepsilon$ where the error term ε has mean 0 and is uncorrelated with $g(U)$ and $\beta = \text{cov}(f(U), g(U)) / \text{var}(g(U))$. Therefore, taking expectations on both sides and reorganising the terms, $E(f(U)) = f(U) - \beta[g(U) - E(g(U))]$. The Monte-Carlo estimator

$$\frac{1}{n} \sum_{i=1}^n \{f(U_i) - \beta[g(U_i) - E(g(U_i))]\}$$

is an improved control variate estimator, equivalent to the one discussed above in the case $\beta = 1$. Determine how much better this estimator is than the basic control variate case $\beta = 1$ by performing simulations. Show that the variance is reduced by a factor of approximately 60. Is there much additional improvement if we use a more general quadratic function of x ?

In this case the optimal value of β is given by $\text{cov}(f(U), g(U)) / \text{var}(g(U)) = .0822 / .0836 = .98325$ and $\text{var}(f(U) - .98325U) = .0013$ resulting in an efficiency gain of around .0822/.0013=63. If we allow a general quadratic function the efficiency gain is even greater than the 2600 observed in problem 86.

87. Ross 6.1&6.2 *Can use OneservQ and replace servicet.*
88. Ross 6.3
89. Ross 6.5&6.6
90. Ross 6.11
91. Ross 6.14
92. Ross 6.15
93. Jobs arrive at a single server queue according to a nonhomogeneous Poisson process with rate function initially 4 per hour, increasing steadily until it reaches 19 per hour after 5 hours and then decreases steadily to 4 per hour after an additional 5 hours. Suppose that the service times are exponential with a mean service time of 1/25 hours. Suppose that whenever the server finds the system empty he goes on a break for 10 minutes (1/6 hours). If, on returning, there are no jobs waiting, he goes on another break. Use simulation to estimate the expected amount of time spent on breaks in the course of a 10 hour day. Assume that the system begins the day empty (so there is an immediate 10 minute break).
94. Show that a single server $M/M/1/\infty/\infty$ queue has steady state distribution for the number in the system which is geometric and find the expected number in the system. Run a simulation and compare the theoretical distribution with the observed amounts of time that the various states are occupied for values of $\frac{\lambda}{\mu} = 0.1, 0.5, 0.9$.
95. In the case $f(x) = \frac{e^x - 1}{e - 1}$, use importance sampling with a probability density function proportional to $g(x) = x$ to integrate over $[0, 1]$. What is the efficiency gain over crude Monte Carlo?
- A suitable probability density function is $g(x) = 2x, 0 < x < 1$. Then $E(f(U)) = E[f(X)/g(X)]$ where $X = U^{1/2}$ has probability density function $g(x)$ so the estimator is the average of sample values $f(U^{1/2})/g(U^{1/2})$ and the variance of the importance sampling estimator if we do n simulations is $\frac{1}{n} \text{var}(f(U^{1/2})/g(U^{1/2})) = \frac{1}{n} \times .0028$. The efficiency gain is about .0822/.0028 or around 30.
96. Suppose I have three different simulation estimators Y_1, Y_2, Y_3 whose means depend on two unknown parameters θ_1, θ_2 . In particular, suppose Y_1, Y_2, Y_3 , are unbiased estimators of $\theta_1, \theta_1 + \theta_2, \theta_2$ respectively. Let us assume for the moment that $\text{var}(Y_i) = 1$, $\text{cov}(Y_i, Y_j) = -1/2$. I want to estimate the parameter θ_1 . Should I use only the estimator Y_1 which is the unbiased estimator of θ_1 , or some linear combination of Y_1, Y_2, Y_3 ? Compare the number of simulations necessary for a certain degree of accuracy (say we wish the variance of the estimator to be less than 1/10000). Simulate 10,000 multivariate normal random variables (Y_1, Y_2, Y_3) with this covariance matrix and $\theta_1 = 1, \theta_2 = 2$ and compare your suggested estimator with Y_1 in this case. (Note: If X is a

column vector of 3 independent standard normal random variables, μ a column vector and A a 3×3 matrix of constants then $\mu + AX$ is multivariate normal with mean vector μ and covariance matrix AA' .

Suppose the covariance matrix of (Y_1, Y_2, Y_3) is denoted by V . We wish to find a linear combination

$$a_1 Y_1 + a_2 Y_2 + a_3 Y_3$$

such that the expected value is θ_1 and the variance is minimized. Since the expected value is θ_1 , $E(a_1 Y_1 + a_2 Y_2 + a_3 Y_3) = a_1 \theta_1 + a_2(\theta_1 + \theta_2) + a_3 \theta_2 = \theta_1$ and so $a_2 = -a_3$ and $a_1 + a_2 = 1$. Therefore the combination takes the form $(1-a)Y_1 + aY_2 - aY_3$. The variance of the estimator is $AV A'$ where $A = [1-a \ a \ -a]$. This gives

$$(1-a)^2 + 2a^2 - a(1-a) + a(1-a) + a^2 = 1 - 2a + 4a^2$$

this variance is minimum and equals $3/4$ when $a = \frac{1}{4}$. The simple estimator Y_1 has variance 1. If I want the variance less than $1/10000$ I need to solve for n , $\frac{3/4}{n} \leq \frac{1}{10000}$ giving $n \geq 7500$.

97. In the case $f(x) = \frac{e^x - 1}{e - 1}$, use $g(x) = x$ as a control variate to integrate over $[0, 1]$. Find the optimal linear combination using estimators $\frac{1}{2}[f(U/2) + f(1 - U/2)]$ and $\frac{1}{4}[f(U/2) + f(\frac{1}{2} - U/2) + f(\frac{1}{2} + U/2) + f(1 - U/2)]$, an importance sampling estimator and the control variate estimator above. What is the efficiency gain over crude Monte-Carlo?

A modification of the matlab function optimal will do. Our X matrix is of the form $X = [Y1' Y2' Y3' Y4']$ and $cov(X)$ gives the covariance matrix V . Then with $Z = \text{ones}(1, 4)$, $V1 = \text{inv}(V)$, we find $b = V1 * Z / (Z' * V1 * Z) = [.366 \ .622 \ -.217 \ .229]$ and the combination

$$.366 Y_1 + .622 Y_2 - .217 Y_3 + .229 Y_4$$

has efficiency gain in excess of 60,000.

98. For independent uniform random numbers U_1, U_2, \dots define the random variable $N = \text{minimum}\{n; \sum_{i=1}^n U_i > 1\}$.

Estimate $E(N)$ by crude Monte Carlo simulation and using antithetic random numbers. What is the efficiency of the use of antithetic random numbers. Can you suggest any other variance reduction techniques? What do you think is the value of $E(N)$?

Define the functions $h_n(x) = P[N > n] = P[U_1 + \dots + U_n \leq x]$ when $x \leq 1$. Then check that $h_1(x) = x$ and

$$h_n(x) = \int_0^x h_{n-1}(x-u) du$$

and so $h_n(1) = 1/n!$. Therefore

$$E(N) = \sum_{n=0}^{\infty} P(N > n) = \sum_{n=0}^{\infty} \frac{1}{n!} = e$$

99. We are interested in the equilibrium number of customers in an $M/G/1/\infty/\infty$ system with Poisson(.8) arrivals and service time distribution $\text{Gamma}(1.1,1)$. This distribution has probability density function

$$f(x) = \frac{1}{\Gamma(1.1)} x^{1/10} e^{-x}, x > 0.$$

Conduct a crude Monte Carlo simulation designed to determine the probability of 0,1,2,... customers in the system.

Repeat using a suitable $M/M/1/\infty/\infty$ queue as a control variate. What is the efficiency gain due to using the control variate?

100. Ross 8.1 Suppose we want to estimate the parameter

$$\theta = \int_0^1 e^{x^2} dx.$$

Show that generating a random uniform[0,1] U and then using the estimator

$$\frac{1}{2} e^{U^2} (1 + e^{1-2U})$$

is better than generating two independent uniform U_1, U_2 and using

$$\frac{1}{2} (\exp(U_1^2) + \exp(U_2^2)).$$

Notice that the antithetic variate estimator is

$$\frac{1}{2} (e^{U^2} + e^{(1-U)^2}) = \frac{1}{2} e^{U^2} (1 + e^{1-2U})$$

and so this estimator uses antithetic random numbers. Since they are applied to the function e^x which is monotone, it results in a variance reduction. $\text{var}(\exp(u.^2))$ gives .2255 and » $\text{var}(.5*(\exp(u.^2)+\exp((1-u).^2)))$ gives .0279 so the efficiency gain is around .2255/.0279 or about 8.

101. Ross 8.2 Explain how to use antithetic random numbers to estimate by simulation the value of the integral

$$\int_0^1 \int_0^1 \exp((x+y)^2) dy dx$$

and compare the use of antithetic to crude Monte Carlo.

We wish to use antithetic random numbers to integrate $\int_0^1 \int_0^1 e^{(x+y)^2} dy dx = 4.8992$. If we generate two random variables X and Y both uniform on $[0, 1]$ we would like a large value of $(X + Y)^2$ to be balanced by a subsequent small value. A simple use of antithetic random numbers takes the form

$$\frac{1}{2} [\exp\{(U_1 + U_2)^2\} + \exp\{(2 - U_1 - U_2)^2\}]$$

To determine the extent of the variance reduction, this has variance around 11.4 while a crude estimator using only one function evaluation has variance around 36.2. Therefore after adjusting for the number of function evaluations the efficiency gain due to the use of antithetic random numbers is $18.1/11.4$ or around 1.6. Very little improvement due to using antithetic random numbers in this case.

102. Ross 8.3 Let $X_i, i = 1, \dots, 5$ be independent exponential random variables each with mean 1. Define

$$\theta = P\left[\sum_{i=1}^5 X_i > 21.6\right].$$

- (a) Explain how to use simulation to estimate θ .
- (b) Give an estimator based on antithetic random variates
- (c) What is the efficiency of the antithetic method?

*Note in this case that the distribution of X_i is $\text{Gamma}(1, 1)$. The sum can be generated in MATLAB using `sum(-(1:5).*log(rand(1,5)))`. So the estimator is $I(\sum -i \ln(U_i) > 21.6)$. Be careful when using antithetic random numbers to average the estimators itself not the random variables that the estimator uses (in this case the exponential random variables). So the antithetic estimator is*

$$.5[I(\sum -i \ln(U_i) > 21.6) + I(\sum -i \ln(1 - U_i) > 21.6)]$$

The following MATLAB code `c=[]; a=[];`

for `i=1:10000`

*`u=rand(1,5); c=[c (sum(-(1:5).*log(u))>21.6)];`*

`a=[a .5((sum(-(1:5).*log(u))>21.6)+(sum(-(1:5).*log(1-u))>21.6))];`*

end

`[mean(c) mean(a)]`

`[var(c) var(a)]` results in variances 0.1437 for crude and 0.0582 for antithetic which, after adjusting for the number of function evaluations, implies an efficiency of $\frac{.1437}{2 \times .0582} : 1.2345$ which again gives very little improvement over crude Monte Carlo.

103. Ross 8.4 Show that if X, Y have the same distribution then

$$\text{var}\left(\frac{X+Y}{2}\right) \leq \text{var}(X)$$

Use this to show that a pair of antithetic random numbers will never have a larger variance than using a single random number. (However it may be worse than using two independent random numbers)

Here

$$\begin{aligned} \text{var}\left(\frac{X+Y}{2}\right) &= \frac{1}{4}(\text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y)) \\ &= \frac{1}{2}(\text{var}(X) + \text{cov}(X, Y)) \leq \text{var}(X) \end{aligned}$$

since $\text{cov}(X, Y) \leq \sqrt{\text{var}(X)}\sqrt{\text{var}(Y)} = \text{var}(X)$. The variance using two independent random numbers is $\text{var}(X)/2$. Using two antithetic random numbers to generate X and Y , $\text{cov}(X, Y) \leq 0$ and so

$$\text{var}\left(\frac{X+Y}{2}\right) = \frac{1}{2}(\text{var}(X) + \text{cov}(X, Y)) \leq \text{var}(X)/2$$

(better than two independent random numbers).

104. Ross 8.5 If Z is a standard normal random variable, use antithetic random numbers to estimate

$$\theta = E[Z^3 e^Z].$$

Design your simulation above so that a 95 percent confidence interval for the value of θ has a length no greater than 0.1.

Here are three possible estimators. There are of course many others as well. We could use antithetic random numbers: (this is the first estimator $\hat{\theta}_1$), or a control variate $g(Z) = Z^3$ since we know that $E(Z^3) = 0$ for $\hat{\theta}_2$, the second estimator or finally we could use importance sampling from a normal $N(\mu, 1)$ probability density function for the third estimator where μ is a parameter chosen roughly to minimize the variance of the estimator. For $\hat{\theta}_1$, recall that use of antithetic random numbers to generate a Normal(0,1) variate is equivalent using the pair $Z, -Z$. This is because if Φ denotes the standard normal c.d.f. and if $Z = \Phi^{-1}(U)$ for some uniform[0,1] random variable then it is easy to show that $-Z = \Phi^{-1}(1 - U)$. The three suggested estimators are therefore

$$\begin{aligned} \hat{\theta}_1 &= \frac{1}{2}(Z^3 e^Z + (-Z)^3 e^{-Z}) = \frac{1}{2}Z^3(e^Z - e^{-Z}) \\ \hat{\theta}_2 &= Z^3(e^Z - 1) \\ \hat{\theta}_3 &= X^3 e^X \frac{\exp(-X^2/2)}{\exp(-(X-\mu)^2/2)} \text{ where } X = \mu + Z \end{aligned}$$

To determine the estimator and its variance in Matlab

```
z=randn(1,100000);      est=.5*(z.^3).*(exp(z)-exp(-z));
mean(est);   (ans = 6.7509)
var(est) (ans = 2.2750e+003)
```

In order that the 95% confidence interval has length less than or equal to 0.1 we require that

$$4S/\sqrt{n} \leq 0.1$$

where $S = 47.69$ and this gives $n \geq (40S)^2$ and this is around 3.64 million. This is the number of antithetic pairs required so the total number of function evaluations is twice this.

105. Ross 8.10

106. Ross 8.6 & 8.9

- (a) Give a pair of exponential random variables with mean 2 which are negatively correlated. Repeat for a pair of exponential random variables with positive correlation.
- (b) Explain why it is unusual to achieve much additional variance reduction from the use of antithetic variates if we use a control variate and then introduce an antithetic variate.

107. Ross 8.12 Show that the variance of a weighted average

$$\text{var}(\alpha X + (1 - \alpha)W)$$

is minimized over α when

$$\alpha = \frac{\text{var}(W) - \text{cov}(X, W)}{\text{var}(W) + \text{var}(X) - 2\text{cov}(X, W)}$$

Determine the resulting minimum variance. What if the random variables X, W are independent?

If we regress I on X we obtain

$$c = -\frac{\text{cov}(X, I)}{\text{var}(X)}$$

If X is $U[0,1]$, then $E(IX) = a$, $a < 1$ and $\text{cov}(I, X) = E(IX) - E(I)E(X) = a^2/2 - a/2 = -a(1-a)/2$. $\text{var}(X) = 1/12$ so $c = (a(1-a)/2)/(1/12) = 6a(1-a)$. The control variate estimator is

$$I + 6a(1-a)(X - \frac{1}{2}).$$

How much variance reduction results? For a crude estimator, $\text{var}(I) = a(1-a)$. For the control variate estimator, $\text{var}(I + 6a(1-a)(X - \frac{1}{2})) = 1 - R^2 = a(1-a)[1 - 3a(1-a)]$ where $R^2 = (\text{cov}(X, I))^2/(\text{var}(X)\text{var}(I))$. Therefore the maximum possible gain in efficiency occurs when $a(1-a)$ is maximized, i.e. $a = 1/2$, and in this case the ratio of the two variances is given by $[1 - 3a(1-a)]^{-1} = 4$. (b) if X is exponential, then $E(IX) = \int_0^a xe^{-x}dx = 1 - (1+a)e^{-a}$ and $\text{cov}(I, X) = E(IX) - E(I)E(X) = 1 - (1+a)e^{-a} - (1 - e^{-a}) = -ae^{-a}$, $\text{var}(X) = 1$ so $c = ae^{-a}$. The control variate estimator is

$$I + ae^{-a}(X - 1).$$

How much variance reduction results? For a crude estimator, $\text{var}(I) = e^{-a}(1 - e^{-a})$. For the control variate estimator, $\text{var}(I + ae^{-a}(X-1)) = e^{-a}(1 - e^{-a}) - a^2e^{-2a}$. The ratio of the two variances is given by $e^{-a}(1 - e^{-a}) / (e^{-a}(1 - e^{-a}) - a^2e^{-2a})$. (c) We know that I and X are negatively correlated because I is a non-increasing function of X (it jumps DOWN as X increases past a).

108. Ross 8.17 Suppose we arrange five elements initially in a random order e.g. (5,3,4,1,2). At each stage one of the elements is randomly selected and put at the front of the list so that element i is chosen with probability p_i where $p_i = i/15$, for $i = 1, 2, 3, 4, 5$. For example if the element 2 is chosen and we started with the configuration above,
109. Ross 8.18
110. Ross 8.19(=8.20 in third edition) The strategy is to generate points (X, Y) independent in the rectangle with base 1 and height b and then record $I = 1$ if $Y < g(X)$ and 0 otherwise. The expected value of I is

$$\theta = E(I) = \frac{\text{area under } g}{\text{area of rectangle}} = \frac{\int g(x)dx}{b}.$$

and $\text{var}(I) = \theta(1 - \theta)$ since I is a 0-1 random variable. Note that

$$\begin{aligned} \text{var}(bI) &= E(\text{var}(bI|X)) + \text{var}(E(bI|X)) \\ &\geq \text{var}(E(bI|X)) \\ &= b^2 \text{var}(g(X)/b) \\ &= \text{var}(g(X)), \quad X \text{ is } U[0, 1] \end{aligned}$$

111. Ross 8.20
112. Ross 8.22 We would generate independent normal X and Y and average the values of e^{XY} . If we were to use the Box-Muller generator for $X = 1 + Z_1$ and $Y = 1 + Z_2$, then $Z_1 Z_2 = (-2 \ln(U_1)) \cos(2\pi U_2) \sin(2\pi U_2)$. One might try as a control variate $g(X, Y) = e^X$ or alternatively e^Y . This can be used since $E(e^X) = e^{3/2}$. Alternatively since $e^x = 1 + x + x^2/2 + \dots$ we might try as control variate a quadratic function of XY like $g(XY) = \beta_0 + \beta_1(XY) + \beta_2(XY)^2$. (d) If we wish to balance large values of XY with small ones, then using antithetic random numbers simultaneously for both X and Y is equivalent to using the estimator

$$\frac{1}{2}(e^{XY} + e^{(2-X)(2-Y)}).$$

Does this decrease the variance over using pairs of independent observations? Note that $\text{var}(e^{XY})$ is REALLY LARGE e.g. 8×10^8 while $\text{var}(\frac{1}{2}(e^{XY} + e^{(2-X)(2-Y)}))$ seems a bit smaller indicating a small improvement due to antithetic random numbers. (e) It is possible to improve on the estimator since the random variable XY given $X = x$ has a conditional normal distribution with mean x and variance x^2 .

Therefore $E(e^{XY}|X) = e^{X+X^2/2}$ and this is an unbiased estimator with smaller variance than e^{XY} . This helps us indentify the true value.

$$\int_{-\infty}^{\infty} e^{x+x^2/2} \frac{1}{\sqrt{2\pi}} e^{-(x-1)^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{2x-1/2} dx$$

and since $e^{2x} \rightarrow \infty$ as $x \rightarrow \infty$, this is an integral that does not even converge (i.e. it is infinite). No wonder we were having trouble with variance reduction. We are chasing infinity. In fact the larger your sample, the larger will the apparent value of the monte carlo integral be. (g) In general, we might hope to improve on the estimator $e^X + X^2/2$ by using a control variate that is a quadratic function of X , say $\beta_0 + \beta_1 X + \beta_2 X^2$. This will have known expected value $\beta_0 + \beta_1 + 2\beta_2$. The coefficients β_i may be estimated by a preliminary regression. However, there is no hope if the integral we are trying to approximate is infinite.

113. Ross 8.26

114. Ross 8.27

115. Ross 8.28

116. Ross 8.32

117. Ross 8.33 A system experiences shocks with a Poisson rate of 1 per hour. The initial damage associated with each shock is a random variable with probability density function $f(x) = xe^{-x}$, $x > 0$ and after a period of time s following the shock the residual damage is $xe^{-\alpha s}$. The system fails when the total damage at a given time exceeds C . We wish to estimate the probability p that the system fails by time t . Explain how to use importance sampling to efficiently estimate p in the case that p is small.

118. Ross 8.35

- (a) This result can be shown counting the repetitions of each name.
- (b) Use that $E\{Y|N(x)\} = P\{Y = 1|N(x)\} = 1/N(x)$.
- (c) Use the equalities $E\{Y\} = E\{E\{Y|N(x)\}\}$ and $E\{1/N(i)\} = E\{1/N(j)\}$.
- (d) W is a mirror image of Y , so that the argument holds as in (b).
- (e) W and Y are negatively correlated which reduces the variance of $n(W + Y)/2$.

119. Suggest three different Monte Carlo methods for approximating the integral $E(Z^3 e^Z)$ where Z is Normal(0,1). Find the optimal linear combination of these estimators and determine the efficiency of this optimal linear combination compared with a crude Monte Carlo estimator. Notice that the function e^Z is

close to 1 when Z is close to 0 so we might use as a control variate the function $g(Z) = Z^3$ where $E(g(Z)) = 0$. The control variate estimator is then an average of values of the form

$$Z^3 e^Z + cZ^3$$

where $c = -\text{cov}(Z^3, Z^3 e^Z) / \text{var}(Z^3)$. Since the function $Z^3 e^Z$ is monotone we may also use antithetic random numbers and average the values of

$$\frac{1}{2}(Z^3 e^Z - Z^3 e^{-Z})$$

Also since it appears that large values of Z are more important than small ones to the integral one might use importance sampling, generating X from a $N(\mu, 1)$ distribution (call the corresponding density function $g_\mu(x)$) and then averaging

$$X^3 e^X \frac{g_0(X)}{g_\mu(X)}.$$

We could then choose μ so that the variance of this estimator is, as nearly as possible, minimized. Note that

$$X^3 e^X \frac{g_0(X)}{g_\mu(X)} = X^3 e^{X(1-\mu)} e^{-\mu^2/2}.$$

120. Suppose $f(X)$ and $g(Y)$ are two **increasing** functions of X and Y respectively, where X and Y are both exponentially distributed random variables with rate parameters 2 and 3 respectively. How would you (as efficiently as possible) use a simulation method to estimate

$$E[f(X)] - E[g(Y)]$$

where we assume that you have a $U[0, 1]$ random number generator with output U_1, U_2, \dots . Taking into account that $X_i = -(1/2) \ln(U_i)$ is an exponential with rate 2 and $Y_i = -(1/3) \ln(U_i)$ is an exponential with rate 3, we can use the same random numbers to generate both expectations. This approach will reduce the variance because

$$\text{Var}\{f(X) - g(X)\} = \text{Var}\{f(X)\} + \text{Var}\{g(X)\} - 2\text{Cov}\{f(X), g(X)\}$$

where the covariance will be positive.

121. Define

$$f(x) = x^{10} e^{-x}$$

and suppose we wish to estimate the integral

$$\theta = \int_0^1 f(x) dx$$

using Monte Carlo integration. Give three methods of using simulation to determine θ as efficiently as possible.

- (a) Implement all three of these methods, giving the estimator of θ using at most 1000 evaluations of the function $f(x)$. Which of your estimators is closest to the true value 0.036?
- (b) Determine the optimal linear combination of these three estimators and its efficiency compared with a crude Monte Carlo estimator.
122. Shocks to a system arrive at times according to a Poisson process with parameter 1 (per hour). Whenever there are five or more shocks within one hour the system fails. Explain how to simulate the probability p that the system fails within 10 hours using

- (a) A crude monte Carlo simulation
- (b) Importance sampling designed to reduce the variance of the estimator. *For importance sampling we wish more failures for a more efficient estimator and this can be achieved by increasing the intensity of the Poisson process. i.e. run the simulation with intensity of shocks $\lambda > 1$. Let I be the indicator of the event that the system fails. Then compute the average of the values*

$$I \frac{f_1(N)}{f_\lambda(N)}$$

with $N =$ the number of shocks in the interval $[0, 10]$ This is the only thing that varies when we change the intensity since given the value of N the shocks themselves are uniformly distributed on the interval $[0, 10]$ and

$$f_\lambda(x) = \frac{(10\lambda)^x e^{-10\lambda}}{x!}, x = 0, 1, \dots$$

Note that this becomes

$$I \frac{\exp(10(\lambda - 1))}{\lambda^N}$$

Alternatively we could multiply by the ratio of the joint probability density function of N and the N spacings Z_1, Z_2, \dots, Z_N between the shocks under the Poisson process with parameter λ . Note that since the Z_i are exponential(λ) random variables, this corresponds to averaging the values

$$\begin{aligned} I \frac{\prod_{i=1}^N \exp(-Z_i)}{\prod_{i=1}^N \{\lambda \exp(-\lambda Z_i)\}} &= I \{\lambda^{-N} \prod_{i=1}^N \exp((\lambda - 1)Z_i)\} \\ &= I \{\lambda^{-N} \exp((\lambda - 1) \sum_{i=1}^N Z_i)\} \end{aligned}$$

and this is almost the same thing as above.

123. If U_1 and U_2 are independent uniform $[0, 1]$ variates, compare the the expected values and variances for the following estimators. *Explain.*
- (a) $T_1 = \frac{1}{2}[e^{U_1} + e^{U_2}]$ This is a crude estimator with sample size 2

- (b) $T_2 = \frac{1}{2}[e^{U_1} + e^{1-U_1}]$ This is an antithetic estimator. Variance better than crude since e^x is monotone so $\text{var}(T_2) < \text{var}(T_1)$
- (c) $T_3 = \frac{1}{2}(e^{U_1/2} + e^{(1+U_2)/2})$ This is a stratified sample with one observation on $[0, .5]$ and another on $[.5, 1]$. Because the stratified sample estimator has variance that is smaller than the within strata variances and since the function is monotone, these are less than the variance over the whole interval, $\text{var}(T_3) < \text{var}(T_1)$.

124. Define

$$f(x) = x^{-1/3}e^{-x}$$

and suppose we wish to estimate the integral

$$\theta = \int_0^2 f(x)dx$$

using Monte Carlo integration. Give estimators based on (i) a control variate (ii) a stratified sample (iii) importance sampling to determine θ as efficiently as possible.

- (a) Determine the efficiency of each of these methods compared with crude Monte Carlo. (i) Notice that the term $x^{-1/3}$ varies between ∞ and $2^{-1/2}$ over the range of this integral but the term e^{-x} is relatively more stable. Therefore we might try as a control variate the function $g(x) = x^{-1/3}$ where $\int_0^2 x^{-1/3}dx = \frac{3}{2}2^{2/3}$. Now $\theta = 2E(f(2U))$ So average values of

$$\hat{\theta}_{CV} = \frac{2}{n} \sum_{i=1}^n (f(2U_i) - g(2U_i)) + \frac{3}{2}2^{2/3}$$

with efficiency given by

$$\text{var}(f(2U_i)) / \text{var}(f(2U_i) - g(2U_i))$$

- (ii) Suppose we stratified the sample with $n/2 = m$ drawn from each of the intervals $[0, 1]$ and $[1, 2]$ (note that this is not an optimal allocation, however). Then

$$\hat{\theta}_{st} = \frac{1}{m} \sum_{i=1}^m f(U_i) + \frac{1}{m} \sum_{i=m+1}^n f(1 + U_i)$$

having efficiency

$$\frac{2\text{var}(f(2U_i))}{\text{var}(f(U_i)) + \text{var}(f(1 + U_i))}$$

- (iii) Suppose we generate X from probability density function $g(x) = cx^{-1/3}, 0 < x < 2$ where $c = 2^{1/3}3^{-1}$. X has cumulative distribution function $G(x) = 2^{-2/3}x^{2/3}, 0 < x < 2$. Then $G^{-1}(U) = 2U^{3/2}$ and

$$c^{-1}E(e^{-X}) = c^{-1} \int_0^2 e^{-x} g(x)dx = \int_0^2 e^{-x} x^{-1/3}dx = \theta$$

so the importance sample estimator is

$$\hat{\theta}_{IM} = \frac{1}{cn} \sum_{i=1}^n e^{-X_i}, \quad \text{where } X_i = 2U_i^{3/2}$$

having efficiency

$$\frac{2c^2 \text{var}(f(2U_i))}{\text{var}(e^{-X_i})}$$

- (b) Find the optimal linear combination of these three estimators and determine its efficiency. For the Monte Carlo integration we make the substitution $y = x/2$ to obtain function $h(y) = 2(2y)^{-1/3}e^{-2y}$ which should be integrated over the interval $[0, 1]$ and the usual Monte Carlo approach can be applied. (i) For the control variate method we can use $g(x) = e^{-x}$. (ii) It turns out that the efficient intervals for the stratified sampling are $(0, 1.1)$ and $(1.2, 2)$. (iii) In the case of importance sampling one can use the probability density function $g(x) = cx^{-1/3}$, where c is approximately .420. (a) The most efficient method in this case is the control variate approach. (b) As the last method is less efficient than crude Monte Carlo method, the best linear combination should be determined using the formula in the notes.

125. Suppose you are given independent $U[0, 1]$ random variables U_1, U_2, \dots . Give a formula or algorithm for generating a single random variable X where X has the

- (a) probability density function $f(x) = \frac{3x^2}{2} + x, 0 < x < 1$. Use composition
- (b) probability (mass) function $f(x) = (1-p)^{x-1}p, x = 1, 2, \dots, 0 < p < 1$. This is a geometric distribution-consult notes.
- (c) probability density function $f(x) = \frac{1.2976}{x^2+x+3}, 0 < x < \infty$, otherwise $f(x) = 0$. You may use acceptance rejection. $f(x) \leq \frac{1.2976}{(x+1/2)^2} = cg(x)$ where $g(x) = \frac{1}{2}(x+1/2)^{-2}$ and $c = 2.59$. Generate from $g(x)$ using inverse transform.
- (d) probability (mass) function

$$P[X = i] = \binom{10}{i} (1/3)^i (2/3)^{10-i}, \quad i = 0, 1, \dots, 10$$

This is a binomial-see notes and text. You may either add 10 Bernoulli variables e.g. $\sum_{i=1}^{10} I(U_i \leq 1/3)$ or one of the other methods.

126. Suppose $U_i, i = 1, 2, 3, \dots, n$ are independent uniform $[0, 1]$ random variables variates and $f(X)$ and $g(Y)$ are two **increasing** functions of X and Y respectively, where X and Y are both exponentially distributed random variables with rate parameters 2 and 3 respectively. How would you (as efficiently as possible) use a simulation method to estimate

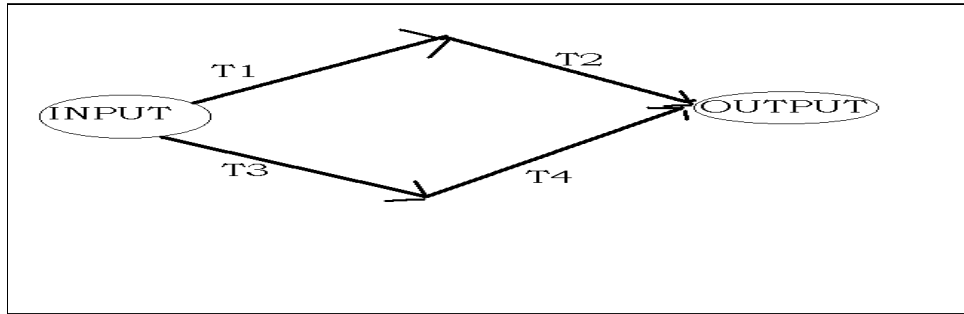


Figure 6:

- (a) $E[f(X)] - E[g(Y)]$ Use common random numbers to generate X and Y i.e. $X = F_X^{-1}(U)$ and $Y = F_Y^{-1}(U)$ or alternatively $X = -\frac{1}{2} \ln(U)$ and $Y = -\frac{1}{3} \ln(U)$.
- (b) $E[f(X)]$ Since f is monotone there is a known benefit from use of antithetic random numbers. e.g. $X_1 = -\frac{1}{2} \ln(U)$ and $X_2 = -\frac{1}{2} \ln(1 - U)$
- (c) $E[f(X)] + E[g(Y)]$ May generate X and Y using antithetic random numbers. e.g. $X = -\frac{1}{2} \ln(U)$ and $Y = -\frac{1}{3} \ln(1 - U)$.

For each Explain how to measure the efficiency with respect to a crude Monte Carlo simulation.

127. Consider the following network of four components (components are represented by directed arcs).The network operates as long as there is flow permitted from the input (source) to the output (sink). The arrows indicate the direction of flow possible. Suppose that the lifetimes of the component i is given by the random variables $T_i, i = 1, 2, \dots, 4$.

- (a) Write an expression involving the random variables $T_i, i = 1, 2, \dots, 4$. for the lifetime of the system.
- (b) Explain how to simulate the expected lifetime of the system if the components have independent lifetimes with an exponential distribution with expected value 2. Provide one such simulation using the uniform random numbers below

U	.92	.51	.24	.43	.35
$-\ln(U)$.08	.67	1.43	.84	1.05

- (c) The first 10 simulations were placed in a vector T of length 10. We then ran the following in MATLAB,

```

mean(T)
ans = 1.17
var(T)
ans = 0.41

```

Determine how large a simulation would be required to provide the average system lifetime within .01 with confidence 95% (i.e. the length of a 95% confidence interval is 0.02).

128. The following data was obtained on the times of individuals in a queueing system which has a Poisson arrivals process.

0, 4, 2, 0, 5, 3, 4, 7, 0, 2, 4, 3, 6, 1, 0, 2, 4, 7, 6, 3, 8, 0. We wish to plan a simulation study in order to estimate the average time spent by a customer in the system when the system is in steady state.

- (a) Give an estimator of this average based on this data.
- (b) Construct a 95% confidence interval for the average time spent by a customer in the system.

129. Ross 9.1

130. Ross 9.2 To determine whether a die was fair 1000 rolls of the die were recorded with the result that the number of times the die landed i , $i=1, \dots, 6$ was given by 158, 172, 164, 181, 160, 165 respectively. Using the chi-squared test determine whether the die was fair and compare the results using (a) the chi-squared distribution approximation and (b) a simulated p -value. (a) For the chi-squared approximation use the last formula on p. 190 in the text.

$$T = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i}$$

and the observed value is $T = 2.18$. Since $n = 6$, compare with a chi-squared with 5 degrees of freedom. Since

$$P[\chi_5^2 > 2.18] = .8236 > 0.05$$

there is no evidence to reject the hypothesis that the model fits. This should be simulated as well. The following code can be used. This gives an approximate p -value of around 0.82 quite consistent with the chi-squared approximation and also leading to accepting the hypothesis that the die is fair.

```
chs=[];
for i=1:50000; u=rand(1,1000);
N=[sum(U<=1/6) sum((U>1/6)&(U<=2/6)) sum((U>2/6)&(U<=3/6)) sum((U>3/6)&(U<=4/6))
sum((U>4/6)&(U<=5/6)) sum(U>5/6)];
chs=[chs sum((N-1000/6).^2./(1000/6))];
end
p=mean(chs>=2.18);
```


131. Ross 9.3 Approximate the p -value for the hypothesis that the following 10 values are uniform[0,1] random numbers. [.12, .18, .06, .33, .72, .83, .36, .27, .77, .74]
Here using the Kolmogorov Smirnov test

$$D = \max\left\{\frac{j}{10} - U_{(j)}, U_{(j)} - \frac{j-1}{10}\right\} = 0.24$$

To obtain the significance probability determine

$$P[D > 0.24]$$

when data is generated from the assumed model.

132. Ross 9.4 Again using the K-S test,

$$D = \frac{9}{67}$$

133. Ross 9.5 Use Kolmogorov-Smirnov test.

$$D = .392254$$

134. Ross 9.6. Approximate the p value of the test that the following data come from a binomial distribution with parameters $(8, p)$ where p is unknown. $X = [6, 7, 3, 4, 7, 2, 6, 3, 7, 8, 2, 1, 3, 5, 8, 7]$.

(the estimated value for p is .6172 and the corresponding binomial probabilities for $x = 0, 1, 2, \dots, 8$ are 0.0005 0.0059 0.0336 0.1082 0.2181 0.2813 0.2268 0.1045 0.0211). The expected numbers in the 9 cells are $16 * \text{probabilities}$ and since the first few of these are very small (i.e. cells corresponding to 0,1,2,3) they might be pooled as should the last two cells 7&8. So the expected numbers after pooling are

	≤ 3	4	5	6	7&8	
Observed Number	6	1	1	2	6	and
Expected Number	2.37	3.49	4.5	3.62	2.01	

$$\chi^2 = \frac{(6 - 2.37)^2}{2.37} + \dots + \frac{(6 - 2.01)^2}{2.01} = 18.7$$

We can estimate the p value by repeating these steps: (1) Generate 16 values X_i from $\text{Bin}(16, .6172)$. (2) Estimate from the sample the value of $p^* = \frac{1}{16} \sum_{i=1}^{16} X_i$. (3) Compute the χ^2 statistic as before but using the $\text{bin}(8, p)$ probabilities (4) Determine whether this statistic is > 18.7 . (5) Repeat steps 1-4 m times to obtain the observed fraction of times the $\chi^2 > 18.7$. This is the approximate p -value of the test if m is large.

135. Ross 9.7 For the K-S test,

$$D = .5717$$

136. Ross 9.8

- (a) i. Generate $n + 1$ uniform random variables
 ii. Generate $n + 1$ exponential random variables
 iii. The sum of the exponentials is the $n + 1$ 'st arrival (event) of a Poisson process.
 iv. Given that the $n + 1$ 'st event of a Poisson process occurs at time τ , the first n event times are uniformly distributed on the ordered values of n uniform $[0, \tau]$. i.e. the event times of the Poisson process Y_j are related to exponential random variables X_i according to:

$$\begin{aligned} Y_0 &= 0 \\ Y_1 &= Y_0 + cX_1, \dots \\ Y_{n+1} &= Y_n + cX_{n+1} \\ \text{where } c &= \frac{1}{\sum_{i=1}^{n+1} X_i} \end{aligned}$$

137. Ross 9.9 Again we may use the K-S test.
 138. Show that if Q_{ij} is a transition matrix of a Markov chain satisfying $Q_{ij} = Q_{ji}$ for all i, j , and if we define a new Markov chain transition matrix by

$$\begin{aligned} P_{ij} &= Q_{ij} \min\left(\frac{b(j)}{b(i)}, 1\right), j \neq i \\ P_{ii} &= 1 - \sum_{j \neq i} P_{ij} \end{aligned}$$

for some non-negative numbers $b(i)$, then the probabilities $\pi_i = b(i) / \sum_j b(j)$ forms a stationary distribution for the Markov chain P .

139. Simulate a Markov Chain on the values $\{1, 2, 3, \dots, 20\}$ such that the chain only moves to neighbouring states (e.g. from 4 to 3 or 5) and the stationary distribution of the chain is $\pi_i = i/210, i = 1, 2, \dots, 20$. For simplicity we assume that the states 1. and 20 are neighbours. Define $b(i) = i$ for $i = 1, 2, \dots, 20$. In this case we can use a chain with transition probability matrix

$$\begin{aligned} P_{ij} &= \frac{1}{2} \min\left(\frac{b(j)}{b(i)}, 1\right), j = i - 1 \text{ or } i + 1, \text{ and } i = 2, \dots, 19 \\ P_{1,20} &= P_{20,1} = \frac{1}{2} \\ P_{ii} &= 1 - \sum_{j \neq i} P_{ij} \end{aligned}$$

Such a chain will have stationary distribution proportional to $b(i)$.

140. Little's formula is a formula relating the arrival rates into a general queuing system, the average number in the system and the average waiting time of individuals in the system. If $\lambda = \text{Poisson arrival rate}$, $L = \text{expected number of}$

customers in the system and W = *average* waiting time of a customer, then Little's formula is

$$L = \lambda W.$$

Verify Little's formula by simulation for a $M/G/1/\infty/\infty$ queue with arrival rate $\lambda = 2$ and with service time distribution Gamma(2,1/6); i.e. probability density function

$$f(x) = 36xe^{-6x}, x > 0.$$

141. Consider an $M/M/1/10/\infty$ queue. This corresponds to the usual one-server queue with Poisson arrivals and exponential service times but with a finite capacity of 10. Whenever there are 10 customers in the system, any arrival "balks", or leaves without joining the system. Find the equilibrium distribution of the number in the system in steady state and compare this with the number obtained by simulating the queue when $\lambda = 3$ and $\mu = 4$.
142. For a general $M/G/1/\infty/\infty$ there is no simple formula giving the steady-state distribution of the number of customers in the system (except in the special case that the service distribution is exponential). However there is a very useful result allowing us to calculate the average number of customers in the queue in steady state. This is called the **Pollaczek-Khintchine** formula.

$$L_q = \frac{\lambda^2 \sigma^2 + \rho^2}{2(1 - \rho)}$$

where ρ = (expected service time)/(expected interarrival time) = $\lambda \times$ (expected service time), σ^2 = *variance* of service time. What is the expected number in the queue if the service times are (a) a constant value c . (b) Exponential with mean c . (c) uniform[0,2c]. Check the above formula by simulation in case $\lambda = 1, c = .9$. Which of the above service times leads to the queue of smallest average length? For (a) $\sigma^2 = 0$ and $\rho = \lambda c$ so $L_q = \frac{\rho^2}{2(1-\rho)}$. For (b), $\sigma^2 = c^2$ and $L_q = \frac{\lambda^2 c^2 + \rho^2}{2(1-\rho)}$. For (c), the variance of a uniform is $c^2/3$ and $L_q = \frac{\lambda^2 c^2/3 + \rho^2}{2(1-\rho)}$. The shortest queue length corresponds to the constant service time (in all cases the expected service time remains the same). In general, for fixed expected service time, the smaller the variance of service time, the shorter the queue on average.

143. Consider an $M/G/1/\infty/\infty$ queue with the following type of service distribution. Customers require one of two types of service in proportions $p, 1 - p$ and the resulting service times are exponentially distributed with parameters μ_1, μ_2 respectively. Let X be the service time of a random customer. Show that $E(X) = p/\mu_1 + (1 - p)/\mu_2$ and $var(X) = 2p/\mu_1^2 + 2(1 - p)/\mu_2^2 - (E(X))^2$. Simulate this queue for values of the parameters $\lambda = 1, p = .5$, and $(\mu_1, \mu_2) = (1, \frac{2}{3}), (\frac{1}{2}, 2)$. Use these simulations to check the above Pollaczek-Khintchine formula. Which values of the parameters seems to result in the shortest queue? Explain.

144. Arrivals to a self-serve gasoline pump occur according to a Poisson process with arrival rate $\lambda = 1/2$ per minute. There are 4 pumps and the service times (in minutes) appear to have probability density function $f(x) = \frac{3}{4}x^2(2-x)$, $0 < x < 2$. Experience shows that only one car will wait for the next available pump. i.e. if there are more than 5 cars in the system, the next arrival will leave immediately. If each customer that is served produces an average profit of \$1.00, conduct a simulation designed to show whether another pump should be added. Assume that the lease of an additional pump will cost the gas station owner a total of 100 dollars per day, that the station remains open for a total of 16 hours during the day.
145. An analyst has developed a queueing model for a parking lot. The lot has space for N cars, and during the daytime shift, cars arrive at the parking lot following a Poisson process with rate or intensity function $10 - ct$ cars per hour at time t where $0 < t < 8$, c some constant. The length of time that a car spends in the lot is assumed to be exponential with mean 1 hour.
- (a) Find the equilibrium distribution of the number of cars in the lot **in the case** $c = 0$.
- (b) Assuming that a lot with N spaces costs \$9 per hour to maintain, (ignore initial costs of building the lot) and the cars pay \$1 per hour to park, what is the expected net profit from operating the lot when N is very large?
- (c) Simulate this system **in the case** $c = 1$ using the uniform(0,1) random numbers
- 0.20 0.60 0.21 0.75 0.93 0.42 0.53 0.67 0.02 0.38
 0.25 0.27 0.02 0.45 0.48 0.85 0.29 0.84 0.68 0.83
- until there have been at least two departures. Give a table like the following

Clock Time	System State	Future Event List

146. Suppose below that you are able to generate independent $U[0, 1]$ random variables U_1, U_2, \dots, U_n . Give a formula or algorithm for generating a single random variable X having the probability density function $f(x) = \frac{4x^3}{2} + \frac{3}{2}x^2$, $0 < x < 1$. In this case the c.d.f. is

$$F(x) = \frac{1}{2}F_1(x) + \frac{1}{2}F_2(x)$$

where

$$\begin{aligned} F_1(x) &= x^4, 0 < x < 1 \\ F_2(x) &= x^3, 0 < x < 1 \end{aligned}$$

so we can use the composition method. Generate U_1, U_2 independent $U[0, 1]$ and put

$$\begin{aligned} X &= U_2^{1/4} \text{ if } U_1 < 1/2 \\ &= U_2^{1/3} \text{ if } U_1 \geq 1/2 \end{aligned}$$

147. If U is a uniform random variate on $[0, 1]$, find a function of U which has probability density function $f(x) = 1/x^2$, $x \geq 1$, and elsewhere $f(x) = 0$. This is simple inverse transform $F^{-1}(U) = 1/(1 - U)$ or alternatively $1/U$.
148. Give an algorithm for generating events according to a nonhomogeneous Poisson process with intensity function $\lambda(t) = \begin{cases} .5 & \text{for } 0 < t < 1 \\ 1 & \text{for } 1 < t < 2 \end{cases}$ and generate points using the uniform random numbers and their logarithms below

	U	$-\ln(U)$
	0.35	1.05
(a)	0.74	0.30
	0.35	1.05
	0.12	2.12
	0.62	0.44

- i. $t=0$
 - ii. $t = t - \ln(U_1)$ where U_1 is $U[0, 1]$
 - iii. generate U_2 with $U[0, 1]$ distribution and go to (ii) if $U_2 > \lambda(t)$ (note this can only happen if $t < 1$).
 - iv. else, register an event at t and go to (ii)
 - v. if $T > T$, stop
149. If U_1 and U_2 are independent uniform $[0, 1]$ variates, which of the following statistics has smaller variance? *Explain.*
- (a) i. $T_1 = -\frac{1}{2}[\log U_1 + \log U_2]$
 - ii. $T_2 = -\frac{1}{2}[\log(U_1) + \log(1 - U_1)]$

For part i, this is the average of two INDEPENDENT exponential random variables each with parameters 1. For (ii) we have a similar average only the exponential random variables are obtained using antithetic random numbers. Therefore both have the same mean but the second average using antithetic random numbers has smaller variance.

150. Give a formula or algorithm for generating a single random variable with probability density function $f(x) = \frac{k}{x^2+3x+1}$, $0 < x < \infty$, $k = 1.16168$, otherwise $f(x) = 0$.

In this case it is clear that

$$f(x) \leq \frac{k}{x^2 + 2x + 1} = \frac{k}{(x+1)^2} = kg(x), \text{ say}$$

Here we can easily generate from the density function $g(x) = 1/(1+x)^2$ by inverse transform. $G(x) = 1 - 1/(1+x)$ and

$$G^{-1}(U) = \frac{1}{1-U} - 1.$$

Therefore we generate from $f(x)$ by acceptance-rejection:

$$\text{Let } X = \frac{1}{1-U_1} - 1$$

If $U_2 \leq \frac{X^2+2X+1}{X^2+3X+1}$ then accept X , otherwise return to step 1.

151. In estimating the expected value $E(g(X))$ of a complicated function of an exponential random variable. X and Y were generated using antithetic random variables and two U of W co-op students disagreed over which estimator was better

$$(i) \quad g\left(\frac{X+Y}{2}\right) \quad \text{or} \quad (ii) \quad \frac{g(X)+g(Y)}{2}.$$

Which estimator do you prefer? Explain with reference to the mean and variance. *This question is a deliberate trap for the unwary. The second estimator is an average of the values of the function and satisfies*

$$E\left[\frac{g(X)+g(Y)}{2}\right] = E[g(X)]$$

since X and Y have the same distribution. However the first estimator does not have the correct expected value, except in the rare case that the function g is linear. In other words

$$E\left(g\left(\frac{X+Y}{2}\right)\right) \neq E(g(X))$$

If we don't even have the right expected value there is no point in discussing its variance!

152. Explain what the following MATLAB code is designed to do:

```
(a) mu=r*T-sigma^2*T/2;
    z=norminv(rand(1,100000),mu,sigma*sqrt(T));
    ST1=S0*exp(z); ST2=S0*exp(2*mu-z);
    v=exp(-r*T)*0.5*(max(ST1-K,0)+max(ST2-K,0));
    mean(v)
```

You are supposed to note first that Z is a vector of $N(rT - \frac{\sigma^2 T}{2}, \sigma^2 T)$ and so $ST1 = S_0 \exp(Z)$ has the distribution of the stock price at time T when the initial stock price is S_0 . The random variable $2\mu - Z$ has the same distribution as Z but is obtained from Z using antithetic random numbers. So the quantity

$$v = \frac{1}{2}(e^{-rT}[(ST1 - K)^+ + (ST2 - K)^+])$$

is the discounted to present return from two call options, one antithetic to the other. This is therefore an estimator of the current price of a call option with exercise price K using antithetic random numbers.

153. Suppose we wish to estimate $E(X)$ for $X \sim EXP(1)$ using an estimator of the form $\hat{\theta} = \frac{X+Y}{2}$ where X and Y have the same $EXP(1)$ distribution but may be dependent.

(a) Show that

$$var(\hat{\theta}) \leq var(X).$$

(b) Determine $var(\hat{\theta})$ in the case of

- i. X, Y independent
- ii. X, Y antithetic
- iii. X, Y use common random numbers. (Hint: $\int_0^1 \ln(u) \ln(1-u) du = 0.355$ and $\int_0^1 (\ln(u))^2 du = 2$).
- iv. What is the efficiency of the antithetic and common random numbers methods compared with crude Monte Carlo.

Part (a) is a simple calculation since

$$\begin{aligned} var(\hat{\theta}) &= \frac{1}{4}(var(X) + var(Y) + 2cov(X, Y)) \\ &= \frac{1}{2}(var(X) + cov(X, Y)) \\ &\leq \frac{1}{2}(var(X) + \sqrt{var(X)}\sqrt{var(Y)}) \\ &\leq var(X) \text{ since } var(X) = var(Y). \end{aligned}$$

$cov(X, Y) \leq \sqrt{var(X)}\sqrt{var(Y)}$ since a correlation coefficient is always less than or equal to one. When X, Y are independent we have

$$\begin{aligned} var(\hat{\theta}) &= \frac{1}{4}(var(X) + var(Y) + 2cov(X, Y)) \\ &= \frac{1}{2}(var(X) + cov(X, Y)) \\ &= \frac{1}{2}(var(X)) = 1/2 \end{aligned}$$

When they are antithetic, $cov(X, Y) = E((- \ln(U))(- \ln(1-U))) - 1 = .355 - 1 = -.645$

$$\begin{aligned} var(\hat{\theta}) &= \frac{1}{4}(var(X) + var(Y) + 2cov(X, Y)) \\ &= \frac{1}{2}(var(X) + cov(X, Y)) \\ &= \frac{1}{2}(1 - .645) \\ &= 0.1775 \end{aligned}$$

Using common random numbers, $cov(X, Y) = var(X) = 1$.

$$\begin{aligned} var(\hat{\theta}) &= \frac{1}{4}(var(X) + var(Y) + 2cov(X, Y)) \\ &= \frac{1}{2}(var(X) + cov(X, Y)) \\ &= 1 \end{aligned}$$

The relative efficiency using antithetic random numbers is $(1/2)/0.1775$ and using common random numbers is $1/2$.

154. A given random variable represents the stresses applied to a bridge in a given day. We wish to estimate by simulation a (very small) probability of the form

$$\theta = P(T > a)$$

where a represents a critical level of stress resulting in damage to the structure. The stress T is a function of several variables including $X = \text{windspeed}$. It is assumed that X has a $Gamma(2, 5)$ distribution with mean 10,

- (a) Find the control variate estimator of the form

$$I(T > a) - \beta(X - EX)$$

- (b) using the following MATLAB output obtained from simulated values of X and T both stored in vectors;

```
I=(T>a) % generates vector of indicator function
of failures
mean(X)= 9.5,      mean(I)=.01
cov([X' I']) % sample covariance matrix of X and
I
50    .6
.6    .0099
```

- (c) What is the efficiency of this control variate estimator compared with crude Monte Carlo?
- (d) Give an importance sampling estimator designed to estimate θ more efficiently.

The control variate estimator requires that we estimate the coefficient β using

$$\hat{\beta} = \frac{cov(I(T > a), X)}{var(X)} = \frac{0.6}{50} = .012$$

Therefore the estimator is the average of the terms

$$I(T > a) - 0.012(X - 10)$$

and this average is

$$0.01 - 0.012(9.5 - 10) = 0.016$$

The efficiency is the ratio of the variances and

$$\begin{aligned} \text{var}(I(T > a) - \hat{\beta}(X - EX)) &= \text{var}(I(T > a)) - \hat{\beta}^2 \text{var}(X) \\ &= 0.0099 - (.012)^2(50) = .0027 \end{aligned}$$

The variance of the crude is .0099 and so the efficiency is (.0099)/(.0027)=3.67. An importance sampling estimator that will improve efficiency is to generate X using a Gamma(2, b) distribution with $b > 5$. We then average the values of

$$I(T(X) > a) \frac{f_5(X)}{f_b(X)}$$

where

$$f_b(x) = \frac{1}{b^2} x e^{-x/b}$$

155. A system fails either if

- (a) Component 1 or 2 fail and
Component 4 or 5 fail and
Component 2 or 3 or 4 fail and
Components 1 or 3 or 5 fail.

Suppose that the lifetimes of the component i is given by the random variables $T_i, i = 1, 2, \dots, 5$.

- (b) Write an expression involving the random variables $T_i, i = 1, 2, \dots, 5$ for the lifetime of the system.
- (c) Explain how to simulate the expected lifetime of the system if the components have independent lifetimes with an exponential distribution with expected value 2. Provide one such simulation using the uniform random numbers below

U	.92	.51	.24	.43	.35
$-\ln(U)$.08	.67	1.43	.84	1.05

- (d) The first 100 simulations were placed in a vector T of length 100. We then ran the following in MATLAB,

`mean(T) = 1.17`

`var(T) = 0.41`

Determine how large a simulation would be required to provide the average system lifetime within .01 with confidence 95% (i.e. the length of a 95% confidence interval is 0.02).

The lifetime of the system is the random variable

$$\max(\min(T_1, T_2), \min(T_4, T_5), \min(T_4, T_3, T_2), \min(T_1, T_3, T_5))$$

Based on the uniform random numbers provided we would simulate the values of T_i as, respectively $T_i = .16, 1.34, 2.86, 1.68, 2.10$ and the simulated lifetime of the system as

$$\max(.16, 1.68, 1.34, .16) = 1.68$$

If we want the average lifetime within .01 we want the width of a 95% confidence interval

$$2\sigma/\sqrt{n} \leq 0.01$$

and solving this we obtain

$$n \geq 16,400$$

Corrections to be made:

Question 47 code is for one day but picture is for one month.

69 There is a better choice of c and p available-see Cheng's algorithm.