# SUPPLEMENTARY PROBLEMS CHAPTER 3

1. Prove: *If $X$ is a random variable uniform on the integers $\{0, \ldots, m-1\}$ and if $Y$ is any integer-valued random variable independent of $X$, then the random variable $W = (X+Y) \bmod m$ is uniform on the integers $\{0, \ldots, m-1\}$.*

2. Consider a multiplicative generator of the form $x_{n+1} = ax_n (\text{mod m})$. Prove that if $m = m_1 m_2$ is the product of relative primes (i.e. $\gcd(m_1, m_2) = 1$), then the period of the generator $x_{n+1} = ax_n (\text{mod } m)$ is the least common multiple of the generators with moduli $m_1$ and $m_2$.

3. Verify that for the serial correlation statistic $C_{j,}$

$$var(C_j) = \begin{cases} 4/45n & \text{for} \quad j = 0 \\ \frac{7}{72n} & \text{for} \quad j \geq 2, \text{ even} \\ \frac{13}{144n} & \text{for} \quad j \geq 1, \text{ odd} \end{cases}$$

4. Consider the turbo-pascal generator $x_{n+1} = (134775813x_n + 1) \bmod 2^{32}$. Generate a sequence of length 5000 and apply the serial correlation test. Is there evidence of dependence?

5. Consider the multiplicative pseudo-random number generator

$$x_{n+1} = ax_n \bmod 150$$

starting with seed $x_0 = 7$. Try various values of the multiplier $a$ and determine for which values the period of the generator appears to be maximal.

6. Briefly indicate an efficient algorithm for generating one random variable from each of the following distributions

   (a) a random variable $X$ with $U[-1, 2]$ probability density function.

   (b) a random variable $X$ with probability density function. $f(x) = \frac{3}{16}x^{1/2}, 0 < x < 4$

   (c) A discrete random number $X$ having probability function $P[X = x] = (1-p)^x p, x = 0, 1, 2, ..., p = 0.3$.

   (d) A random variable $X$ with the normal distribution, mean 1 and *variance* 4.

   (e) A random variable $X$ with probability density function

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

   for constant $c = 1/(2 - 5e^{-1})$.

   (f) 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 |

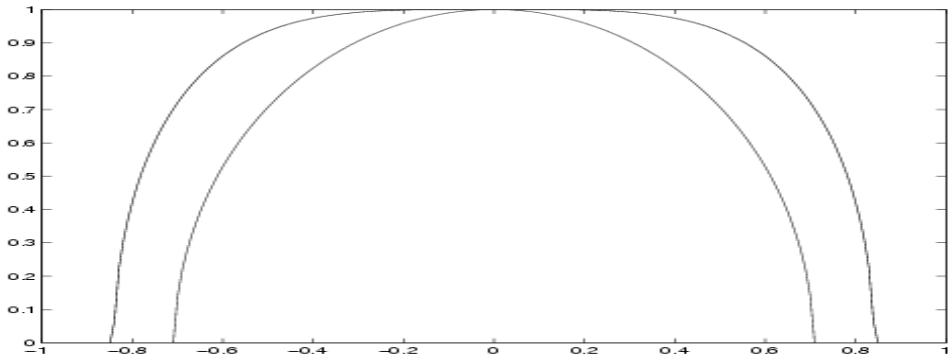


Figure 1: The region $\{(x,y); -1 < x < 1,\ \ y > 0,\ \sqrt{1-2x^2} < y < \sqrt{1-2x^4}\}$

7. Evaluate the following integral by simulation:

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

8. 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.

9. What is the probability density function of $X = a(1-\sqrt{U})$ where $U \sim U[0,1]$?

10. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm for generating observations from a distribution with probability density function

$$f(x) = \frac{(x-1)^3}{4}$$

for $1 \le x \le 3$. Record the time necessary to generate the sample mean of 5,000 random variables with this distribution.

11. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm for generating observations from a distribution with probability density function $\frac{(x-20)}{200}$ for $20 \le x \le 40$. Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. .

12. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm 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.

13. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm for generating observations from a discrete distribution with $P[X = j] = (2/3)(1/3)^j; j = 0, 1, ....$Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. .

14. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm for generating observations from a distribution with probability density function $f(x) = e^{-x}, 0 \leq x < \infty$. Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. Compute as well the sample variance and compare withe 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?

15. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm for generating observations from a distribution which has probability density function $f(x) = x^3, 0 < x < \sqrt{2}$. 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?

16. Assume that you have available a Uniform[0,1] random number generator. Give a precise algorithm for generating observations from a discrete distribution with probability function

| | | | | | | |
|---|---|---|---|---|---|---|
| x= | 0 | 1 | 2 | 3 | 4 | 5 |
| P[X=x]= | 0.1 | 0.2 | 0.25 | 0.3 | 0.1 | 0.05 |

Record the time necessary to generate the sample mean of 100,000 random variables with this distribution. Compare the sample mean and variance with their theoretical values. How large would the simulation need to be if we wanted to estimate the mean within 0.01 with a 95% confidence interval? How much difference does it make to the time if we use an optimal binary search? If we use the alias method?

17. Give an algorithm for generating a random variable with probability density function

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

Discuss the efficiency of your approach.

18. The interarrival times between consecutive buses at a certain bus stop are independent uniform$[0, 1]$ random variables starting at clock time $t = 0$. You arrive at the bus stop at time $t = 1$. Determine by simulation the expected time that you will have to wait for the next bus. Is it more or less than 1/2 ? Explain.

19. The interarrival times between consecutive buses at a certain bus stop are independent 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 expected value $\frac{1}{2}$. Compare your average wait times with the expected time between buses and explain how your results are possible.

20. Explain how the following algorithm works and what distribution is generated.

    (a) Let $I = 0$
    (b) Generate $U \sim U[0,1]$and set $T = U$.
    (c) Generate $U^*$. IF $U \leq U^*$return $X = I + T$.
    (d) Generate $U$. If $U \leq U^*$go to c.
    (e) $I = I + 1$. Go to b

21. Assume that a option has payoff at expiry one year from now ($T = 1$) given by the function $g(S_T) = 0, S_T < 20$, and $g(S_T) = \frac{S_T - 20}{S_T}, S_T > 20$. What is the approximate present value of the option assuming that the risk-neutral interest rate is 5 percent, the current price of the stock is 20, and the annual volatility is 20 percent. Determine this by simulating 1000 stock prices $S_T$ and averaging the discounted return from a corresponding option. Repeat with 100000 simulations. What can you say about the precision?

22. (*Hedging with futures*). I need to buy 10000 barrels of heating oil on November 1 2004. On June 1, 2004, I go long a December futures contract which allows me to purchase 10000 barrels of heating oil on December 1 for \$40 per barrel. Suppose we have observed that the price of heating oil is lognormally distributed with monthly volatility 3 percent. The spot interest rate is presently 5 percent per annum

    (a) What is the value of the oil future on November 1 as a function of the current price of oil?
    (b) Determine by simulation what is the standard deviation of the value of my portfolio on November 1 assuming I sell the futures contract at that time.
    (c) How much difference would it have made if I had purchased the optimal number of futures rather than 1000?

23. What distribution is generated by the following algorithm where $U$ is uniform$[0,1]$ and $V$ is uniform $[-\sqrt{2/e}, \sqrt{2/e}]$?

    (a) GENERATE $U, V$
    (b) PUT $X = V/U$
    (c) IF $-ln(U) < X^2/4$,GO TO a.; ELSE RETURN $X$.
    *(Hint: First prove the following result due to Kinderman and Monahan, 1977: Suppose $(U, V)$ are uniformly distributed on the region $\{(u,v); 0 \leq v \leq 1,\ 0 \leq u \leq \sqrt{f(\frac{v}{u})}\}$ for some integrable function $1 \geq f(x) \geq 0$. Then $V/U$ has probability density function $cf(x)$ with $c = 1/\int f(x)dx$. )*

24. *(Cox, Ingersoll, Ross model for interest rates)* Use the Milstein approximation to simulate paths from a CIR model of the form

$$dr_t = k(b - r_t) + \sigma\sqrt{r_t}dW_t$$

and plot a histogram of the distribution of $r_1$assuming that $r_0 = .05$ for $b = 0.04$. What are the effects of the parameters $k$ and $b$?

25. Simulate independent random variables from the Normal Inverse Gamma distribution using parameter values so that the expected value, the variance, the skewness and the kurtosis of daily returns are respectively 0, 0.004, 0.5 and 4 respectively. Evaluate an at the money call option with time to maturity 250 days using these simulated values and compare the price of the option with the Black-Scholes price. Repeat with an option whose strike is 10% over the initial stock price and one 10% under.

26. Daily relative losses from a portfolio are assumed to follow the NIG distribution with parameters

$$\alpha = 100, \delta = .02, \beta = 0, \mu = 0.$$

In other words if the portfolio is worth $S_t$ at the end of day $t$, then

$$\frac{S_t - S_{t+1}}{S_t}$$

has the above NIG distribution. Assume daily relative losses are independent of one another. Assume $S_0 = 100000$. Use simulation to determine the weekly 99% VAR, i.e. the value $x$ such that

$$P[S_5 - S_0 \leq x] = 0.99$$

Compare this result with the Value at Risk (VaR) if we replace the NIG distribution with the Normal having the same mean and variance.

# CHAPTER 4

1. Consider the *systematic sample* estimator based on the trapezoidal rule:

$$\hat{\theta} = \frac{1}{n}\sum_{i=0}^{n-1} f(V + i/n), V \sim U[0, \frac{1}{n}]$$

Discuss the bias and variance of this estimator. In the case $f(x) = x^2$, how does it compare with other estimators such as crude Monte Carlo and antithetic random numbers requiring $n$ function evaluations. Are there any disadvantages to its use?

2. For any random variables $X, Y$, prove that $P(X \leq x, Y \leq y) - P(X \leq x)P(Y \leq y) = P(X > x, Y > y) - P(X > x)P(Y > y)$ for all $x, y$.

3. A particular (strange) option pays an amount $V(S) = 1/(1+\exp(S(T)-K))$ at time $T$ for some predetermined price $k$. Discuss what you would use for a control variate and conduct a simulation to determine how it performs, assuming geometric Brownian motion for the stock price, interest rate 5%, annual volatility 20% and various initial stock prices, values of $K$ and $T$.

4. Suppose we wish to generate the partial sum of independent identically distributed summands, $S_n = \sum_{i=1}^{n} X_i$ for

   (a) $S_n$ is generated with $X_i$ having the Normal$(0, \sigma^2)$ distribution

   (b) $S_n^*$ is generated with $X_i$ having a student $t$ distribution with $m = 5$ degrees of freedom.
   What is the maximum possible correlation we can achieve between $S_n$ and $S_n^*$? What is the minimum correlation?

5. (Importance sampling in high dimension) Suppose we wish to evaluate

$$E[h(X_1, ..., X_n)]$$

by Monte Carlo where the random variables $X_i$ are independent with probability density function $f(x)$. Suppose instead we generate random variables $Z_i$ from the probability density function $g(z)$.

   (a) Show

$$E[h(X_1, ..., X_n)] = E[h(Z_1, ..., Z_n)\exp\{\sum_{i=1}^{n} q(Z_i)\}]$$

   for some function $q(Z_i)$ such that

$$E[e^{q(Z_i)}] = 1$$

   but

$$\sum_{i=1}^{n} q(Z_i) \to -\infty$$

   in probability.

(b) Then the optimal choice of importance density $g(z)$ is the one which minimizes

$$E[h^2(Z_1, ..., Z_n) \exp\{2 \sum_{i=1}^{n} q(Z_i)\}].$$

Suppose that $g(z)$ is obtained using an exponential tilt from $f(z)$ so that

$$g(z) = \frac{e^{\lambda z}}{m(\lambda)} f(z)$$

and suppose that $h(X_1, ..., X_n) = h(X_1 + ... + X_n) = h(S)$, say. Then the optimal choice of $\lambda$ satisfies

$$E\{(S - K'(\lambda))h^2(S)e^{-\lambda S}\}\} = 0$$

where $K(\lambda)$ is the cumulant generating function of $S$.

6.

7.

8.

# CHAPTER 5

1. (a) Suppose the low $H$ and the close $C$ of a process over a given interval satisfy

$$P[H \geq h | C = c] = \frac{f(2h-c)}{f(c)}, h \geq \max(c, 0)$$

where $f(x)$ is the probability density function of $C$. Show that the joint p.d.f. of $(H, C)$ is given by

$$f(h, c) = -2f'(2h - c)$$

and the conditional cumulative distribution function of $C$ given $H$ by

$$F(c|h) = \frac{f(2h-c)}{f(h)}, c \leq h. \tag{1}$$

Thus, to generate the value of $C$ conditional on $H = h$, we need to solve for $c$ the equation

$$f(2h - c) = f(h)U, \text{where } U \text{ is uniform[0,1]}$$

(b) Consider the special case where $f(x)$ is a Normal(0,$\sigma^2$) probability density function. How does the conditional probability density function (1) change as the variance $\sigma^2$ increases? What consequence does this have for the conditional expectated value

$$E[C|H = h]?$$

(c) Give the analogous results above for the conditional distribution of the close given the low $L$.