Chapter 6

Characteristic Functions and the Central Limit Theorem

6.1 Characteristic Functions

6.1.1 Transforms and Characteristic Functions.

There are several transforms or generating functions used in mathematics, probability and statistics. In general, they are all integrals of an exponential function, which has the advantage that it converts sums to products. They are all functions defined on $t \in \Re$. In this section we use the notation $i = \sqrt{-1}$. For example;

1. *(Probability) Generating function.* $g(s) = E(s^X)$.
2. *Moment Generating Function.* $m(t) = E[e^{tX}] = \int e^{tx} dF$
3. *Laplace Transform.* $\mathcal{L}(t) = E[e^{-tX}] = \int e^{-tx} dF$
4. *Fourier Transform.* $E[e^{-itX}] = \int e^{-itx} dF$
5. *Characteristic function.* $\varphi_X(t) = E[e^{itX}] = \int e^{itx} dF$

Definition 106 *(*Characteristic Function) *Define the characteristic function of a random variable X or or its cumulative distribution function F_X to be the complex-valued function on $t \in \Re$*

$$\varphi_X(t) = E[e^{itX}] = \int e^{itx} dF = E(cos(tX)) + iE(sin(tX))$$

The main advantage of the characteristic function over transforms such as the Laplace transform, probability generating function or the moment generating function is property (a) below. Because we are integrating a bounded function; $|e^{itx}| = 1$ for all $x, t \in \Re$, the integral exists for any probability distribution.

6.1.2 Properties of Characteristic Function.

(a) φ exists for any distribution for X.

(b) $\varphi(0) = 1$.

(c) $|\varphi(t)| \leq 1$ for all t .

(d) φ is *uniformly continuous* . That is for all $\epsilon > 0$, there exists $\delta > 0$ such that $|\varphi(t) - \varphi(s)| \leq \epsilon$ whenever $|t - s| \leq \delta$.

(e) The characteristic function of $a + bX$ is $e^{iat}\varphi(bt)$.

(f) The characteristic function of $-X$ is the complex conjugate $\bar{\varphi}(t)$.

(g) A characteristic function φ is real valued if and only if the distribution of the corresponding random variable X has a distribution that is symmetric about zero, that is if and only if $P[X > z] = P[X < -z]$ for all $z \geq 0$.

(h) The characteristic function of a convolution $F * G$ is $\varphi_F(t)\varphi_G(t)$.

Proofs.

(a) Note that for each x and t, $|e^{itx}|^2 = sin^2(tx) + cos^2(tx) = 1$ and the constant 1 is integrable. Therefore

$$E|e^{itX}|^2 = 1.$$

It follows that

$$E|e^{itX}| \leq \sqrt{E|e^{itX}|^2} = 1$$

and so the function e^{itx} is integrable.

(b) $e^{itX} = 1$ when $t = 0$. Therefore $\varphi(0) = Ee^0 = 1$.

(c) This is included in the proof (a).

(d) Let $h = s - t$. Assume without loss of generality that $s > t$. Then

$$\begin{aligned} |\varphi(t) - \varphi(s)| &= |Ee^{itX}(e^{ihX} - 1)\ | \\ &\leq E[|e^{itX}(e^{ihX} - 1)|] \\ &\leq E[|e^{itX}||e^{ihX} - 1|] \\ &\leq E[|e^{ihX} - 1|]. \end{aligned}$$

But as $h \to 0$ the function $e^{ihX} - 1$ converges to 0 for each $\omega \in \Omega$ and it is dominated by the constant 2. Therefore, by the Lebesgue Dominated Convergence theorem, $E[|e^{ihX} - 1|] \to 0$ as $h \to 0$. So for a given $\epsilon > 0$, we chan choose h sufficiently small, for example $h = |s - t| \le \delta$ such that $|\varphi(t) - \varphi(s)| \le \epsilon$.

(e) By definition, $Ee^{it(a+bX)} = e^{ita}E[e^{itbX}] = e^{iat}\varphi(bt)$.

(f) Recall that the complex conjugate of $a + bi$ is $a - bi$ and of e^{iz} is e^{-iz} when $a, b,$ and z are real numbers. Then

$$\begin{aligned} E[e^{it(-X)}] &= E[e^{-itX}] = E[cos(tX) + isin(-tX)] \\ &= E[cos(tX) - isin(tX)] = \bar{\varphi}(t). \end{aligned}$$

(g) The distribution of the corresponding random variable X is symmetric if and only if X has the same distribution as does $-X$. This is true if and only if they have the same characteristic function. By properties (f) and the corollary below, this is true if and only if $\varphi(t) = \bar{\varphi}(t)$ which holds if and only if the function $\varphi(t)$ takes on only real values.

(h) Put $H = F{*}G$. When there is a possible ambiguity about the variable over which we are integrating, we occasionally use the notation $\int h(x)F(dx - y)$ to indicate the integral $\int h(x)dK(x)$ where $K(x)$ is the cumulative distribution function given by $K(x) = F(x - y)$. Then

$$\int e^{itx}H(dx) = \int e^{itx} \int F(dx - y)G(dy)$$

$$= \int \int e^{it(z+y)}F(dz)G(dy), \text{with } \ z = x - y,$$

and this is $\varphi_F(t)\varphi_G(t)$.

The major reason for our interest in characteristic functions is that they uniquely describe the distribution. Probabilities of intervals can be recovered from the characteristic function using the following inversion theorem.

Theorem 107 *(*Inversion Formula*). If X has characteristic function $\varphi_X(t)$, then for any interval (a, b),*

$$P[a < X < b] + \frac{P[X = a] + P[X = b]}{2} = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)dt$$

Proof. Consider the integral

$$\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)dt = \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \int_{\Re} e^{itx}F(dx)dt$$

$$= \int_{-T}^{T} \int_{\Re} \frac{e^{it(x-a)} - e^{it(x-b)}}{2\pi it} F(dx)dt = \int_{\Re} \int_{-T}^{T} \frac{e^{it(x-a)} - e^{it(x-b)}}{2\pi it} dtF(dx).$$

Note that for real c we have

$$\int_{-T}^{T} \frac{e^{itc}}{2it} dt = \int_{0}^{T} \frac{sin(tc)}{t} dt$$

and so we obtain from above

$$\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)dt = \int_{\Re} \frac{1}{\pi} \{ \int_{0}^{T} \frac{sin(t(x-a))}{t} dt - \int_{0}^{T} \frac{sin(t(x-b))}{t} dt \} F(dx).$$

But as $T \to \infty$, it is possible to show that the integral (this is known as the sine integral function)

$$\frac{1}{\pi} \int_{0}^{T} \frac{sin(t(x-a))}{t} dt \to \begin{cases} -\frac{1}{2}, & x < a \\ \frac{1}{2}, & x > a \\ 0, & x = a \end{cases}$$

Substituting this above and taking limits through the integral using the Lebesgue Dominated Convergence Theorem, the limit is the integral with respect to $F(dx)$ of the function

$$g(x) = \begin{cases} \frac{1}{2}, & x = a \\ \frac{1}{2}, & x = b \\ 1, & a < x < b \\ 0, & elswhere \end{cases}$$

and this integral equals

$$P[a < X < b] + \frac{P[X = a] + P[X = b]}{2}.$$

■

Corollary 108 *If the characteristic function of two random variables X and Y agree, then X and Y have the same distribution.*

Proof. This follows immediately from the inversion formula above. ■

We have seen that if a sequence of cumulative distribution functions $F_n(x)$ converges pointwise to a limit, the limiting function $F(x)$ is not necessarily a cumulative distribution function. To ensure that it is, we require that the distributions are "tight". Similarly if a sequence of characteristic functions converge for each t, the limit is not necessarily the characteristic function of a probability distribution. However, in this case the tightness of the sequence translates to a very simple condition on the limiting characteristic function.

Theorem 109 *(Continuity Theorem) If X_n has characteristic function $\varphi_n(t)$, then X_n converges weakly if and only if there exists a function $\varphi(t)$ which is continuous at 0 such that $\varphi_n(t) \to \varphi(t)$ for each t.* (Note: In this case φ is the characteristic function of the limiting random variable X.)

Proof. Suppose $X_n \Rightarrow X$. Then since the function e^{itx} is a continuous bounded function of x, then

$$E(e^{itX_n}) \to E(e^{itX}).$$

Conversely, suppose that $\varphi_n(t) \to \varphi(t)$ for each t and φ is a continuous function at $t = 0$. First prove that for all $\epsilon > 0$ there exists a $c < \infty$ such that $P[|X_n| > c] \leq \epsilon$ for all n. This is Problem 14 below. This shows that the sequence of random variables X_n is "tight" in the sense that any subsequence of it contains a further subsequence which converges in distribution to a proper cumulative distribution function. By the first half of the proof, $\varphi(t)$ is the characteristic function of the limit. Thus, since every subsequence has the same limit, $X_n \Rightarrow X$. ■

Example 110 *(Problem 18) Suppose $X_n \sim U[-n, n]$. Then the characteristic function of X_n is $\varphi_n(t) = (\sin tn)/tn$. Does this converge as $n \to \infty$? Is the limit continuous at 0?*

Example 111 *(Problem 19) Suppose $X_1, ...X_n \ldots$ are independent Cauchy distributed random variables with probability density function*

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \Re.$$

Then the sample mean $\bar{X}$ has the same distribution as X_1.

Note: We may use the integral formula

$$\int_0^\infty \frac{\cos(tx)}{b^2 + x^2} dx = \frac{\pi}{2b} e^{-tb}, t \geq 0$$

to obtain the characteristic function of the above Cauchy distribution

$$\varphi(t) = e^{-|t|}.$$

6.1.3 Characteristic function of $N(\mu, \sigma^2)$.

The characteristic function of a random variable with the distribution $N(\mu, \sigma^2)$ is

$$\varphi(t) \;=\; exp\{i\mu t - \frac{\sigma^2 t^2}{2}\}.$$

(Note: Recall that for any real constant c,

$$\int_{-\infty}^{\infty} e^{-(x-c)^2/2} dx = \sqrt{2\pi}.$$

Use the fact that this remains true even if $c = it$).

6.2 The Central Limit Theorem

Our objective is to show that the sum of independent random variables, when standardized, converges in distribution to the standard normal distribution. The proof usually used in undergraduate statistics requires the moment generating function. However, the moment generating function exists only if moments of all orders exist, and so a more general result, requiring only that the random variables have finite mean and variance, needs to use characteristic functions. Two preliminary lemmas are used in the proof.

Lemma 112 *For real* x ,

$$e^{ix}-(1+ix-\frac{x^2}{2})=r(x)$$

where $|r(x)|\le min[x^2,\frac{|x|^3}{6}]$. *Consequently,*

$$\varphi(t) \;=\; 1+itE(X)-\frac{t^2}{2}E(X^2)+o(t^2)$$

where $\frac{o(t^2)}{t^2}\to 0$ *as* $t\to 0$.

Proof. By expanding e^{ix} in a Taylor series with remainder we obtain

$$\frac{e^{ix}-1}{i}=x+i\frac{x^2}{2}+i^2\frac{b_2}{2}$$

where $b_n(x)=\int_0^x(x-s)^n e^{is}ds$, and a crude approximation provides $|b_2|\le\int_0^x s^2ds=x^3/3$. Integration by parts shows that $b_2=\frac{2b_1-x^2}{i}$ and substituting this provides the remaining bound on the error term. ∎

Lemma 113 *For any complex numbers* $w_i,\ z_i$, *if* $|z_i|\le 1,\quad |w_i|\le 1$, *then* $|\prod_i z_i-\prod_i w_i|\le\sum_i|z_i-w_i|$.

Proof. This is proved by induction using the fact that

$$\prod_{i=1}^n z_i-\prod_{i=1}^n w_i=(z_n-w_n)(\prod_{i=1}^{n-1}z_i)+w_n(\prod_{i=1}^{n-1}z_i-\prod_{i=1}^{n-1}w_i)\le|z_n-w_n|+|(\prod_{i=1}^{n-1}z_i-\prod_{i=1}^{n-1}w_i)|.$$

∎

This shows the often used result that

$$(1-\frac{c}{n}+o(\frac{1}{n}))^n-(1-\frac{c}{n})^n\to 0$$

and hence that

$$(1-\frac{c}{n}+o(\frac{1}{n}))^n\to e^{-c}\quad\text{as } n\to\infty.$$

Theorem 114 *(Central Limit Theorem) If X_i are independent identically distributed random variables with $E(X_i) = \mu$, $var(X_i) = \sigma^2$, then*

$$S_n^* = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$$

converges weakly to $N(0,1)$.

Proof. Suppose we denote the characteristic function of $\frac{X_i - \mu}{\sigma}$ by $\varphi(t)$. By Lemma 46, $\varphi(t) = 1 - \frac{t^2}{2} + r(t)$ where $\frac{r(t)}{t^2} \to 0$ as $t \to 0$. Then the characteristic function of S_n^* is

$$\varphi^n(t/\sqrt{n}) = \{1 - \frac{t^2}{2n} + o(t^2/n)\}^n.$$

Note that by Lemma 47,

$$|\{1 - \frac{t^2}{2n} + o(t^2/n)\}^n - (1 - \frac{t^2}{2n})^n| \le n\ o(t^2/n) \to 0$$

and the second term $(1 - \frac{t^2}{2n})^n \to e^{-t^2/2}$. Since this is the characteristic function of the standard normal distribution, it follows that S_n^* converges weakly to the standard normal distribution. ∎

6.3 Problems

1. Find the characteristic function of the normal(0,1) distribution. Prove using characteristic functions that if F is the $N(\mu, \sigma^2)$ c.d.f. then $G(x) = F(\mu + \sigma x)$ is the $N(0,1)$ c.d.f.

2. Let F be a distribution function and define

$$G(x) = 1 - F(-x-)$$

   where $x-$ denotes the limit from the left. Prove that $F * G$ is symmetric.

3. Prove that $F * G = G * F$.

4. Prove using characteristic functions that if $F_n \Rightarrow F$ and $G_n \Rightarrow G$, then $F_n * G_n \Rightarrow F * G$.

5. Prove that convolution is *associative*. That

$$(F * G) * H = F * (G * H).$$

6. Prove that if φ is a characteristic function, so is $|\varphi(t)|^2$.

7. Prove that any characteristic function is non-negative definite:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi(t_i - t_j)z_i\bar{z}_j \ \geq 0$$

for all real $t_1, ...t_n$ and complex $z_1, ..., z_n$.

8. Find the characteristic function of the Laplace distribution with density on $\Re$

$$f(x) = \frac{1}{2}e^{-|x|}. \tag{6.1}$$

What is the characteristic function of $X_1 + X_2$ where X_i are independent with the probability density function (6.1)?

9. (Stable Laws) A family of distributions of importance in financial modelling is the *symmetric stable family.* These are unimodal densities, symmetric about their mode, and roughly similar in shape to the normal or Cauchy distribution (both special cases $\alpha = 2$ or 1 respectively). They are most easily described by their *characteristic function*, which, upon setting location equal to 0 and scale equal to 1 is

$$Ee^{iXt} = e^{-|t|^\alpha}.$$

The parameter $0 < \alpha \leq 2$ indicates what moments exist, for except in the special case $\alpha = 2$ (the normal distribution), moments of order less than α exists while moments of order α or more do not. Of course, for the normal distribution, moments of all orders exist. The probability density function does not have a simple closed form except in the case $\alpha = 1$ (the Cauchy distribution) and $\alpha = 2$ (the Normal distribution) but can, at least theoretically, be determined from the series expansion of the probability density valid in case $1 < \alpha < 2$ (the cases, other than the normal and Cauchy of most interest in applications)

$$\begin{aligned} f_c(x) &= \sum_{k=0}^{\infty}(-1)^k\frac{\Gamma((k+1)/\alpha)}{\pi c\alpha k!}\cos(\frac{k\pi}{c})(\frac{x}{c})^k \\ &= \sum_{n=0}^{\infty}(-1)^n\frac{\Gamma((2n+1)/\alpha)}{\pi c\alpha(2n)!}(\frac{x}{c})^{2n} \end{aligned}$$

(a) Let $X_1, ..., X_n$ be independent random variables all with a symmetric stable (α) distribution. Show that $n^{-1/\alpha}\sum_{i=1}^{n} X_i$ has the same Stable distribution.

(b) Verify that the Characteristic function of the probability density function

$$f_1(x) = \sum_{n=0}^{\infty}(-1)^n\frac{\Gamma((2n+1)/\alpha)}{\pi\alpha(2n)!}x^{2n}$$

is given by $e^{-|t|^\alpha}$.(Hint: the series expansion of $\cos(x)$ is

$$\cos(x) = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!}).$$

10. Let Ω be the unit interval and P the uniform distribution and suppose we express each $\omega \in [0,1]$ in the binary expansion which does not terminate with finitely many terms. If $\omega = .\omega_1\omega_2....$, define $R_n(\omega) = 1$ if $\omega_n = 1$ and otherwise $R_n(\omega) = -1$. These are called the *Rademacher functions*. Prove that they are independent random variables with the same distribution.

11. For the Rademacher functions R_n defined on the unit interval, Borel sets and Lebesgue measure, let

    $Y_1 = R_1/2 + R_3/2^2 + R_6/2^3....$

    $Y_2 = R_2/2 + R_4/2^2 + R_7/2^3 + ...$

    $Y_3 = R_5/2 + R_8/2^2 + R_{12}/2^3 + ...$

    Note that each R_i appears only in the definition of one Y_j. Prove that the Y_i are independent identically distributed and find their distribution.

12. Find the characteristic function of:

    (a) The Binomial distribution

    (b) The Poisson distribution

    (c) The geometric distribution

    Prove that suitably standardized, both the binomial distribution and the Poisson distribution approaches the standard normal distribution as one of the parameters $\to \infty$.

13. (*Families Closed under convolution.*) Show that each of the following families of distributions are closed under convolution. That is suppose X_1, X_2 are independent and have a distribution in the given family. Then show that the distribution of $X = X_1 + X_2$ is also in the family and identify the parameters.

    (a) $Bin(n,p)$, with p fixed.

    (b) Poisson (λ).

    (c) Normal (μ, σ^2).

    (d) Gamma (α, β), with β fixed.

    (e) Chi-squared.

    (f) Negative Binomial, with p fixed.

14. Suppose that a sequence of random variables X_n has characteristic functions $\varphi_n(t) \to \varphi(t)$ for each t and φ is a continuous function at $t = 0$. Prove that the distribution of X_n is tight, i.e. for all $\epsilon > 0$ there exists a $c < \infty$ such that $P[|X_n| > c] \le \epsilon$ for all n.

15. Prove, using the central limit theorem, that

$$\sum_{i=0}^{n} \frac{n^i e^{-n}}{i!} \to \frac{1}{2}, \quad \text{as} \quad n \to \infty.$$

16. (*Negative binomial*) Suppose we decide in advance that we wish a fixed number (k) of successes in a sequence of Bernoulli trials, and sample repeatedly until we obtain exactly this number. Then the number of trials X is random and has probability function

$$f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \quad x = k, k+1, \ldots.$$

Use the central limit theorem to show that this distribution can be approximated by a normal distribution when k is large. Verify the central limit theorem by showing that the characteristic function of the standardized Negative binomial approaches that of the Normal.

17. Consider a random walk built from independent Bernoulli random variables $X_i = 1$ with probability $p = \mu/\sqrt{n}$ and otherwise $X_i = 0$. Define the process

$$B_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} X_i$$

for all $0 \leq t \leq 1$. Find the limiting distribution of $B(t)$ and the limiting joint distribution of $B(s)$, $B(t) - B(s)$ for $0 < s < t < 1$.

18. Suppose $X_n \sim U[-n, n]$. Show that the characteristic function of X_n is $\varphi_n(t) = (\sin tn)/tn$. Does this converge as $n \to \infty$? Is the limit continuous at 0?

19. Suppose $X_1, .\ \ X_n$ are independent Cauchy distributed random variables with probability density function

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \Re.$$

Show that the characteristic function of the Cauchy distribution is

$$\varphi(t) = e^{-|t|}.$$

and verify that the sample mean $\bar{X}$ has the same distribution as does X_1. *Note: We may use the integral formula*

$$\int_0^\infty \frac{\cos(tx)}{b^2 + x^2} dx = \frac{\pi}{2b} e^{-tb}, t \geq 0.$$

20. What distribution corresponds to the following characteristic functions?

(a)
$$\varphi(t) = \exp\{ita - b|t|\}.$$

(b)
$$\varphi(t) = \frac{\sin t}{t}.$$

(c)
$$\frac{2 - 2\cos t}{t^2}$$

21. Let X be a discrete random variable which takes only integer values. Show that $\varphi(t)$ is a periodic function with period 2π.

22. Let X be a random variable and $a \neq 1$. Show that the following conditions are equivalent:

(a)
$$\varphi_X(a) = 1$$

(b) φ_X is periodic with period $|a|$.

(c)
$$P[X \in \{\frac{2\pi j}{a}, j = ..., -2, -1, 0, 1, 2, ...\}] = 1$$

23. Let X be a random variable which is bounded by a finite constant. Prove that
$$E(X^n) = \frac{1}{i^n}\frac{d^n}{dt^n}\varphi_X(0)$$

24. What distribution corresponds to the following characteristic functions?

(a)
$$\varphi_X(t) = \frac{1}{2}e^{-it} + \frac{1}{3} + \frac{1}{6}e^{2it}.$$

(b)
$$\varphi_X(t) = \cos(\frac{t}{2})$$

(c)
$$\varphi_X(t) = \frac{2}{3e^{it} - 1}$$

25. Show that if X correponds to an absolutely continuous distribution with probability density function $f(x)$, then
$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi_X(t)dt$$

26. What distribution corresponds to the following characteristic functions?

(a)

$$\varphi_X(t) = \frac{1}{1+t^2}$$

(b)

$$\varphi_X(t) = \frac{1}{3}e^{it\sqrt{2}} + \frac{2}{3}e^{it\sqrt{3}}$$

(c)

$$\varphi_X(t) = \frac{1}{2}e^{-t^2/2} + \frac{1}{2}e^{it}$$