Chapter 5. Distributions of Functions of Random Variables

- ♣ Distributions of Functions of Random Variables
- ♣ Sampling Distribution Theory
- ♣ Random Functions Associated with Normal Distributions
- ♣ The Central Limit Theorem (CLT)
- ♣ Approximations for Discrete Distributions
- ♣ Limiting Moment-Generating Functions
- ♣ Box-Muller Transformation
- ♣ The Beta, Student's t, and F Distributions

□ *Distributions of Functions of Random Variables*

- We discuss the distributions of functions of one random variable X and the distributions of functions of independently distributed random variables in this Chapter.

Example 1. Let X have the p.d.f. $f(x) = xe^{-x^2/2}, \quad 0 < x < \infty$. Then $Y = X^2$ has *an exponential distribution with mean 2*.

Example 2. The p.d.f. of X is $f(x) = \theta x^{\theta-1}, \quad 0 < x < 1, \quad 0 < \theta < \infty$. Then $Y = -2\theta ln(X)$ has *an exponential distribution with mean 2*.

Example 3. Let X have a *logistic distribution* with p.d.f.

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \quad -\infty < x < \infty$$

Then $Y = \frac{1}{1+e^{-X}}$ has a $U(0,1)$ distribution.

Example 4. Let $X_1 \sim b(m,p)$ and $X_2 \sim b(n,p)$ be independent r.v.'s, then $Y = X_1 + X_2 \sim b(m+n,p)$.

Sampling Distribution Theory

♣ The collection of n independent and identically distributed random variables X_1, X_2, ..., X_n, is called *a random sample of size n from the common distribution*, e.g., $X_j \sim N(0,1),\ 1 \le j \le n$.

♣ Some functions of *a random sample*, called *statistics*, are of interest, for examples, *mean* and *variance*. *Sampling distribution theory* refers to the derivation of distributions for functions of a random sample.

Theorem 1: Let $X_1, X_2, \ldots, X_n$ be n independent r.v.'s with respective means $\{\mu_i\}$ and variances $\{\sigma_i^2\}$, then $Y = \sum_{i=1}^n a_i X_i$ has mean $\mu_Y = \sum_{i=1}^n a_i \mu_i$ and variance $\sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma_i^2$, respectively.

Theorem 2: Let $X_1, X_2, \ldots, X_n$ be n independent r.v.'s with respective moment-generating functions $\{M_i(t)\}$, $1 \le i \le n$, then the moment-generating function of $Y = \sum_{i=1}^n a_i X_i$ is $M_Y(t) = \prod_{i=1}^n M_i(a_i t)$.

Corollary: If $X_1, X_2, \ldots, X_n$ are observations of a random sample from a distribution with moment-generating function $M(t)$, then

(a) $M_Y(t) = \prod_{i=1}^n M(t)$, where $Y = \sum_{i=1}^n X_i$.

(b) $M_{\overline{X}}(t) = \prod_{i=1}^n M(t/n)$, where $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

Example 1: Let $X_i \sim b(k,p)$ be a random sample of size n. Define $Y = \sum_{i=1}^n X_i$, then $M_Y(t) = \prod_{i=1}^n (q + pe^t)^k = (q+pe^t)^{kn}$.

Example 2: Let $X_i \sim Gamma(1,\theta)$ be a random sample of size n. Define $Y = \sum_{i=1}^n X_i$, then $M_Y(t) = \prod_{i=1}^n (1-\theta t)^{-1} = 1/(1-\theta t)^n$.

Exercises:

Random Functions Associated with Normal Distributions

♣ In statistical applications, it is usually assumed that the population from which a *sample* is taken is $N(\mu, \sigma^2)$.

Theorem: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $N(\mu, \sigma^2)$. Define $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, then $\overline{X} \sim N(\mu, \sigma^2/n)$.

Theorem: Let $X_j \sim \chi^2(r_j)$, $1 \leq j \leq n$. If $X_1, X_2, \ldots, X_n$ are independent, then $Y = \sum_{i=1}^n X_i \sim \chi^2(r_1 + r_2 + \ldots + r_n)$.

Theorem: Let $Z_1, Z_2, \ldots, Z_n$ be a random sample of size n from $N(0,1)$, then $W = Z_1^2 + Z_2^2 + \ldots + Z_n^2 \sim \chi^2(n)$.

Corollary: Let $\{X_i' s\}$ be independent random variables from $N(\mu_i, \sigma_i^2)$, respectively, then $W = \sum_{i=1}^n (X_i - \mu_i)^2/\sigma_i^2$ is $\chi^2(n)$.

Theorem: Let $\{X_i' s\}$ be observations of a random sample of size n from $N(\mu, \sigma^2)$. Define $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$, then

(a) $\overline{X}$ and S^2 are independent.

(b) $\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n (X_i - \overline{X})^2/\sigma^2 \sim \chi^2(n-1)$.

Example 1: Let X_1, X_2, X_3, X_4 be a random sample of size 4 from the normal distribution $N(76.4, 383)$. Then

(a) $U = \sum_{i=1}^n (X_i - 76.4)^2/383 \sim \chi^2(4)$, $P(0.711 \leq U \leq 7.779)$ = 0.90 - 0.05 =0.85.

(b) $W = \sum_{i=1}^n (X_i - \overline{X})^2/383 \sim \chi^2(3)$, $P(0.352 \leq W \leq 6.251)$ = 0.90 - 0.05 =0.85.

Theorem: Let $X_i \sim N(\mu_i, \sigma_i^2)$, $1 \leq i \leq n$, be independent. Define $Y = \sum_{i=1}^n a_i X_i$, then $Y \sim N(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2)$.

The Central Limit Theorem

Theorem: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $N(\mu, \sigma^2)$. Define $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, then $\overline{X} \sim N(\mu, \sigma^2/n)$.

Theorem: Let $\overline{X}$ be the mean of a random sample $X_1, X_2, \ldots, X_n$ of size n from a distribution with mean μ and variance σ^2. Define $W_n = (\overline{X} - \mu)/(\sigma/\sqrt{n})$. Then

(a) $W_n = (\sum_{i=1}^n X_i - n\mu)/(\sqrt{n}\sigma)$

(b) $P(W_n \leq w) \approx \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \Phi(w)$.

(c) $W_n \sim N(0, 1)$ as $n \to \infty$.

Example 1: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a $\chi^2(1)$. Define $Y = \sum_{i=1}^n X_i$. Then

(a) $Y \sim \chi^2(n)$.

(b) $(Y - n)/\sqrt{2n} \approx N(0, 1)$.

Example 2: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a $U(0, 1)$. Define $Y = \sum_{i=1}^n X_i$. Then

$(Y - 0.5n)/\sqrt{n/12} \approx N(0, 1)$.

Example 3: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a $Bernoulli(p)$. Define $Y = \sum_{i=1}^n X_i$. Then

(a) $Y \sim b(n, p)$.

(b) $(Y - np)/\sqrt{np(1-p)} \approx N(0, 1)$.

Example 4: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from an exponential distribution with mean θ. Define $Y = \sum_{i=1}^n X_i$. Then

(a) $Y \sim Gamma(n, \theta)$.

(b) $(Y - n\theta)/\sqrt{n\theta^2} \approx N(0, 1)$.

Approximations for Discrete Distributions

♣ Use the normal distribution to approximate probabilities for certain discrete-type distributions.

Example 1: Let $Y \sim b(10, 1/2)$. Then

$$\begin{aligned} P(3 \leq Y < 6) &= P(2.5 \leq Y \leq 5.5) \\ &= \Phi(0.316) - \Phi(-1.581) \\ &= 0.6240 - 0.0570 \\ &= 0.5670(0.5683 \;\; by \;\; Table \; II). \end{aligned} \tag{1}$$

Example 2: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a $Poisson(\lambda)$. Define $Y = \sum_{i=1}^{n} X_i$. Then

$$(Y - n\lambda)/\sqrt{n\lambda} \approx N(0, 1).$$

Example 3: Let $Y \sim Poisson(\lambda = 20)$. Then

$$\begin{aligned} P(16 < Y \leq 21) &= P(16.5 \leq Y \leq 21.5) \\ &= P[(16.5 - 20)/\sqrt{20} \leq (Y - 20)/\sqrt{20} \leq (21.5 - 20)/\sqrt{20})] \\ &= \Phi(0.335) - \Phi(-0.783) = 0.4142 \end{aligned} \tag{2}$$

Limiting Moment-Generating Functions

Theorem: If a sequence of moment-generating functions approaches a certain one, say, $M(t)$, then the limit of the corresponding distribution must be the distribution corresponding to $M(t)$.

Example 1: Let $Y \sim b(50, 0.04)$ and let $\lambda = np = 50 \times 0.04 = 2$. Then

$P(Y \leq 1) = 0.400$

$P(Y \leq 1) \approx 0.406$ by a Poisson approximation.

Simulating Continuous Distributions

Theorem 5.1-2 Let X have the cumulative distribution function (c.d.f.) $F(x)$ of the continuous type that is *strictly increasing* (i.e., $F(t) > F(s)$ if $t > s$) in on the support $a < x < b$. Then the r.v. $Y = F(X)$ has a uniform distribution $U(0,1)$.

Proof Since $F(a) = 0$ and $F(b) = 1$. For $0 < y < 1$, we have

$$\begin{aligned} P(Y \le y) &= P(F(x) \le y) &= P(X \le F^{-1}(y)) \\ &= F(F^{-1}(y)) &= y \end{aligned}$$

Thus, Y has a uniform distribution $U(0,1)$.

- Simulating an exponential distribution $f(x) = \frac{1}{2}e^{-x/2}$, $0 < x < \infty$.

(1) $Y = F(X) = 1 - e^{-X/2} \sim U(0,1)$,

(2) Generate a y from $U(0,1)$ and let $y = 1 - e^{-x/2}$

(3) Then $x = -2 \times \ln(1 - y + \epsilon)$,

(4) Repeat steps (2) and (3) for the sample size you request.

Example 5.2-6: Box-Muller Transformation

Box-Muller Transformation Let $\{X_1, X_2\}$ be a random sample from U(0,1), define

$$Z_1 = \sqrt{-2lnX_1}cos(2\pi X_2) \text{ and } Z_2 = \sqrt{-2lnX_1}sin(2\pi X_2).$$

or, equivalently

$$X_1 = \exp\left(-\frac{Z_1^2+Z_2^2}{2}\right) = e^{-q/2} \text{ and } X_2 = \frac{1}{2\pi}arctan\left(\frac{Z_2}{Z_1}\right),$$

which has the Jacobian

$$J = \begin{vmatrix} -z_1e^{-q/2} & -z_2e^{-q/2} \\ \frac{-z_2}{2\pi(z_1^2+z_2^2)} & \frac{z_1}{2\pi(z_1^2+z_2^2)} \end{vmatrix} = -\frac{1}{2\pi}e^{-q/2}.$$

Since the joint p.d.f. of X_1 and X_2 is $f(x_1, x_2) = 1, \;\; 0 < x_1, \;\; x_2 < 1,$

hence the joint p.d.f. of Z_1 and Z_2 is

$$g(z_1, z_2) = |J_{x_1,x_2}| = \frac{1}{2\pi}exp[-({z_1}^2 + {z_2}^2)/2], \quad -\infty < z_1, \;\; z_2 < \infty.$$

The Beta Distribution

Beta $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $0 < x < 1, \quad \alpha, \beta \in N$

Example Let X_1 and X_2 have independent gamma distribution with parameters α, θ and β, θ, respectively. That is the joint probability density function (p.d.f.) of X_1 and X_2 is

$$f(x_1, x_2) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}} x_1^{\alpha-1} x_2^{\beta-1} Exp\left(-\frac{x_1 + x_2}{\theta}\right), \; 0 < x_1, x_2 < \infty, \quad \alpha, \beta \in N$$

Consider

$$Y_1 = \frac{X_1}{X_1 + X_2}, \quad Y_2 = X_1 + X_2$$

or, equivalently,

$$X_1 = Y_1 Y_2, \quad X_2 = Y_2 - Y_1 Y_2$$

The jacobian is

$$J \;=\; \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1 - y_1 \end{vmatrix} \;=\; y_2(1 - y_1) + y_1 y_2 \;=\; y_2.$$

Thus, the joint p.d.f. of Y_1 and Y_2 is

$$g(y_1, y_2) = y_2 \frac{1}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}} (y_1 y_2)^{\alpha-1} (y_2 - y_1 y_2)^{\beta-1} e^{-y_2/\theta},$$

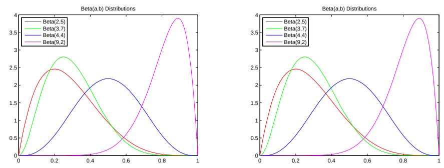
where $0 < y_1 < 1$ and $0 < y_2 < \infty$.

$$g(y_1) = \frac{y_1^{\alpha-1}(1 - y_1)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \frac{y_2^{\alpha+\beta-1}}{\theta^{\alpha+\beta}} e^{-y_2/\theta} dy_2.$$

In particular, when $\theta = 1$, we have a beta distribution

$$g(y_1) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y_1^{\alpha-1}(1 - y_1)^{\beta-1}, \quad 0 < y_1 < 1.$$

What is $E(Y_1)$ and $Var(Y_1)$?



(a) (b)

Figure 1: Beta Distributions.

Student's t and F Distributions

Random variables whose space are intervals or a union of intervals are said to be of the *continuous types*. The *p.d.f.* of a r.v. X of continuous type is an integrable function $f(x)$ satisfying

(a) $f(x) > 0,\ x \in R$

(b) $\int_R f(x)dx = 1$

(c) The probability of the event $X \in A$ is $P(A) = \int_A f(x)dx$

Student's t Let $Z \sim N(0,1)$ and $V \sim \chi^2(r)$ be two independent random variables. Define $T = Z/\sqrt{V/r}$. Then T has a t-distribution with p.d.f.

$$f(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+t^2/r)^{(r+1)/2}}, \quad -\infty < t < \infty$$

F-distribution Let $U \sim \chi^2(r_1)$ and $V \sim \chi^2(r_2)$ be two independent random variables. Define $W = (U/r_1)/(V/r_2)$. Then W has an F-distribution with p.d.f.

$$f(w) = \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{x^{(r_1/2)-1}}{(1+r_1w/r_2)^{(r_1+r_2)/2}}, \quad 0 < w < \infty$$

Proof of the Central Limit Theorem

Theorem: Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $N(\mu, \sigma^2)$. Define $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, then $\overline{X} \sim N(\mu, \sigma^2/n)$.

Theorem: Let $\overline{X}$ be the mean of a random sample $X_1, X_2, \ldots, X_n$ of size n from a distribution with mean μ and variance σ^2. Define $W_n = (\overline{X} - \mu)/(\sigma/\sqrt{n})$. Then

(a) $W_n = (\sum_{i=1}^{n} X_i - n\mu)/(\sqrt{n}\sigma)$

(b) $P(W_n \leq w) \approx \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \Phi(w)$.

(c) $W_n \sim N(0, 1)$ as $n \to \infty$.

(Proof)

$$\begin{aligned} E[\exp(tW_n)] &= E\left\{\exp\left[\left(\tfrac{t}{\sqrt{n}\sigma}\right)\left(\textstyle\sum_{i=1}^{n} X_i - n\mu\right)\right]\right\} \\ &= E\left\{\exp\left[\left(\tfrac{t}{\sqrt{n}}\right)\left(\tfrac{X_1-\mu}{\sigma}\right) + \cdots + \left(\tfrac{t}{\sqrt{n}}\right)\left(\tfrac{X_n-\mu}{\sigma}\right)\right]\right\} \\ &= E\left\{\exp\left[\left(\tfrac{t}{\sqrt{n}}\right)\left(\tfrac{X_1-\mu}{\sigma}\right)\right]\right\} \cdots E\left\{\exp\left[\left(\tfrac{t}{\sqrt{n}}\right)\left(\tfrac{X_n-\mu}{\sigma}\right)\right]\right\}, \end{aligned}$$

which follows from the independence of $X_1, X_2, \cdots, X_n$. Then

$$E[\exp(tW_n)] = \left[M\left(\tfrac{t}{\sqrt{n}}\right)\right]^n, \quad -h < \tfrac{t}{\sqrt{n}} < h,$$

where

$$M(t) = E\left\{\exp\left[t\left(\tfrac{X_i-\mu}{\sigma}\right)\right]\right\}, \quad -h < t < h$$

is the common moment-generating function of each

$$Y_i = \frac{X_i - \mu}{\sigma}, \quad i = 1, 2, \cdots, n.$$

since $E(Y_i) = 0$ and $E(Y_i^2) = 1$, it must be that

$$M(0) = 1, \quad M'(0) = E\left(\frac{X_i - \mu}{\sigma}\right) = 0, \quad M''(0) = E\left[\left(\frac{X_i - \mu}{\sigma}\right)^2\right] = 1$$

Hence, using Taylor's formula with a remainder, we know that there exists a number t_1 between 0 and t such that

$$M(t) = M(0) + M'(0)t + \frac{M''(t_1)t^2}{2} = 1 + \frac{M''(t_1)t^2}{2}.$$

Adding and subtracting $t^2/2$, we have

$$M(t) = 1 + \frac{t^2}{2} + \frac{]M''(t_1) - 1]t^2}{2}.$$

Using this expression of $M(t)$ in $E[\exp(tW_n)]$, we can represent the moment-generating function of W_n by

$$\begin{aligned} E[\exp(tW_n)] &= \left\{1 + \tfrac{1}{2}\left(\tfrac{t}{\sqrt{n}}\right)^2 + \tfrac{1}{2}[M''(t_1) - 1]\left(\tfrac{t}{\sqrt{n}}\right)^2\right\}^n \\ &= \left\{1 + \tfrac{t^2}{2n} + \tfrac{[M''(t_1)-1]t^2}{2n}\right\}^n, \quad -\sqrt{n}h < t < \sqrt{n}h, \end{aligned}$$

where now t_1 is between 0 and $t/\sqrt{n}$. Since $M''(t)$ is continuous at $t = 0$ and $t_1 \to 0$ as $n \to \infty$, we have

$$lim_{n\to\infty}[M''(t_1) - 1] = 1 - 1 = 0$$

Thus,

$$\begin{aligned} lim_{n\to\infty} E[\exp(tW_n)] &= lim_{n\to\infty}\left\{1 + \tfrac{t^2}{2n} + \tfrac{[M''(t_1)-1]t^2}{2n}\right\}^n \\ &= lim_{n\to\infty}\left\{1 + \tfrac{t^2}{2n}\right\}^n = e^{t^2/2} \end{aligned}$$

for all real t. We know that $e^{t^2/2}$ is the moment-generating function of the standard normal distribution, $N(0,1)$. Therefore, the limiting distribution of

$$W_n = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}} \longrightarrow N(0,1) \;\; as \;\; n \to \infty.$$

Order Statistics

□ If $X_1, X_2, \cdots, X_n$ are observations of a random sample of size n from a continuous-type distribution whose p.d.f.is $f(x)$ and c.d.f. is $F(x)$, and let the random variables

$$Y_1 < Y_2 < \cdots < Y_n \quad or \quad X_{(1)} < X_{(2)} < \cdots < X_{(n)} \tag{3}$$

denote the order statistics of this sample, that is,

$$Y_1 \quad is \quad the\ smallest\ of\ X_1, X_2, \cdots, X_n$$

$$\vdots$$

$$Y_r \quad is \quad the\ r-th\ smallest\ of\ X_1, X_2, \cdots, X_n$$

$$\vdots$$

$$Y_n \quad is \quad the\ largest\ of\ X_1, X_2, \cdots, X_n$$

Let f be defined in (a, b) so that $F'(x) = f(x),\ for\ x \in (a, b)$, $0 < F(x) < 1,\ x \in (a, b)$ and $F(a) = 0,\ F(b) = 1$. Then we have

$$\begin{aligned} G_r(y) = P(Y_r \le y) &= \textstyle\sum_{k=r}^{n} \binom{n}{k} [F(y)]^k [1 - F(y)]^{n-k} \\ &= \textstyle\sum_{k=r}^{n-1} \binom{n}{k} [F(y)]^k [1 - F(y)]^{n-k} + [F(y)]^n \end{aligned}$$

Thus the p.d.f. of Y_r could be derived as

$$g_r(y) = G_r{}'(y) = \frac{n!}{(r-1)!(n-r)!}[F(y)]^{r-1}[1 - F(y)]^{n-r} f(y), \quad a < y < b \tag{4}$$

In particular,

$$\begin{aligned} g_1(y) &= n[1 - F(y)]^{n-1} f(y), \quad a < y < b \\ g_n(y) &= n[F(y)]^{n-1} f(y), \qquad a < y < b \end{aligned}$$

(1) If X_i has a U(0,1) distribution, $E(Y_r) = \frac{r}{n+1}$.

(2) If X_j has an exponential distribution with mean 2, $g_1(y) = ne^{-ny},\ y > 0$ and $E(Y_1) = \frac{1}{n}$.