Chapter 5

Distributions of Functions of Random Variables

5.1 Functions of One Random Variable

If two continuous r.v.s X and Y have functional relationship, the **distribution function technique** *and the* **change-of-variable technique** *can be used to find the relationship of their p.d.f.s.*

Thm 5.1. *Let X be a continuous r.v. with p.d.f. $f(x)$.*

1. *If $X = v(Y)$ (where v is injective), the p.d.f. of Y is*

$$g(y) = f(v(y))\,|v'(y)|.$$

2. *If $Y = u(X)$ (where u is surjective), the p.d.f. of Y is*

$$g(y) = \sum_{u(x)=y} \frac{f(x)}{|u'(x)|}.$$

Ex *1.* **loggamma p.d.f.**, *p.215-216 (scanned file)*

Ex *2.* **Cauchy p.d.f.**, *p.216-217 (scanned file)*

Ex *3, p.218 (scanned file)*

Ex 3 comes from the following corollary, a special case of Theorem 5.1.

Cor 5.2. *Suppose $Y \sim U(0,1)$. Let $F(x)$ have the properties of a distribution function (i.e. $F(x)$ is increasing, $F(-\infty) = 0$, and $F(\infty) = 1$). Then the r.v. $X := F^{-1}(Y)$ has the distribution function $F(x)$.*

Ex *4-5, p.219-221 (scanned file)*

The reverse of Corollary 5.2 is as follow.

Cor 5.3. *Let X has the distribution function $F(x)$. Then the r.v. $Y = F(X)$ has the uniform distribution $U(0,1)$.*

Ex *6-8, p.222-223 (scanned file)*

Homework
§*5.1* *1, 3, 5, 7, 13*

Attachment: *Scanned textbook pages of Section 5-1*

5.2 Transformations of Two Random Variables

In one r.v. case, if X had p.d.f. $f(x)$, and $X = v(Y)$, then Y has p.d.f.

$$g(y) = f(v(y))\,|v'(y)|.$$

This **change-of-variable technique** *can be extended to bivariate r.v.s..*

Thm 5.4. *Let X_1 and X_2 be two continuous r.v.s with joint p.d.f. $f(x_1, x_2)$. Suppose that $X_1 = v_1(Y_1, Y_2)$ and $X_2 = v_2(Y_1, Y_2)$. Then the joint p.d.f. of Y_1 and Y_2 is*

$$g(y_1, y_2) = f[v_1(y_1, y_2), v_2(y_1, y_2)]\,|J|,$$

where the **Jacobian** *J is the determinant*

$$J := \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}.$$

Ex *1, p.225-226 (scanned file)*

Ex *2.* **double exponential p.d.f.**, *p.226-227 (scanned file)*

Ex *3.* **beta p.d.f.**, *p.228 (scanned file)*

Homework
§*5.2* *1, 13, 15*

Attachment: *Scanned textbook pages of Section 5-2*

5.3 Several Independent Random Variables

If r.v.s $X_1, \cdots, X_n$ *are* **independent**, *then the joint p.d.f. of* $X_1, \cdots, X_n$ *is the product of the respective p.d.f.s, i.e.* $f_1(x) \cdots f_n(x)$.

If all of the distributions are the same, the $X_1, \cdots, X_n$ *is said to be a* **random sample of size** n **from that common distribution**, *or called* n **independent and identically-distributed (i.i.d.)** *random variables of the common distribution.*

Ex *1-3, p.234-236 (scanned file)*

Here are some properties of n *independent r.v.s.*

Thm 5.5. *Let* $X_1, \cdots, X_n$ *be* n *independent discrete r.v.s where* X_i *has the p.m.f.* $f_i(x)$. *Let the r.v.* $Y = u(X_1, \cdots, X_n)$ *have the p.m.f.* $g(y)$. *Then*

$$E(Y) = \sum_y yg(y) = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} u(x_1, \cdots, x_n) f_1(x_1) \cdots f_n(x_n).$$

Similar result holds for continuous r.v.s.

Thm 5.6. *If* $X_1, \cdots, X_n$ *are independent r.v.s, then* $u_1(X_1), \cdots, u_n(X_n)$ *are also independent r.v.s. In particular, for* $Y = u_1(X_1) \cdots u_n(X_n)$,

$$E(Y) = E[u_1(X_1) \cdots u_n(X_n)] = E[u_1(X_1)] \cdots E[u_n(X_n)].$$

Thm 5.7. *If* $X_1, \cdots, X_n$ *are* n *independent r.v.s with respective means* $\mu_1, \cdots, \mu_n$ *and variances* $\sigma_1^2, \cdots, \sigma_n^2$, *then the mean and the variance of* $Y = \sum_{i=1}^n a_i X_i$, *where* a_i *are constants, are, respectively,*

$$\mu_Y = \sum_{i=1}^n a_i \mu_i \qquad \text{and} \qquad \sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma_i^2.$$

Ex *4-5, p.239 (scanned file)*

Let $X_1, \cdots, X_n$ *be i.i.d.s with mean* μ *and variance* σ^2. *The* **mean of the random sample** $X_1, \cdots, X_n$ *is*

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Then

$$\mu_{\overline{X}} = \sum_{i=1}^n \left(\frac{1}{n}\right) \mu = \mu \qquad \text{and} \qquad \sigma_{\overline{X}}^2 = \sum_{i=1}^n \left(\frac{1}{n}\right)^2 \sigma^2 = \frac{\sigma^2}{n}.$$

The r.v. $\overline{X}$ *has the same mean but has only* $1/n$ *of the variances of the underlying distribution.*

Any function of the sample observations $X_1, \cdots, X_n$ *is called a* **statistic**.

Homework

§*5.3* *1, 3, 9, 11, 17, 23*

Attachment: *Scanned textbook pages of Section 5-3*

5.4 The Moment-Generating Function Technique

Ex *1, p.242 (scanned file)*

Thm 5.8. *If $X_1, \cdots, X_n$ are independent r.v.s with respective m.g.f. $M_{X_i}(t)$, $i = 1, \cdots, n$, then the m.g.f. of $Y = \sum_{i=1}^n a_i X_i$ is*

$$M_Y(t) = \prod_{i=1}^n M_{X_i}(a_i t).$$

Proof.

$$\begin{aligned} M_Y(t) &= E[e^{tY}] = E[e^{ta_1X_1}e^{ta_2X_2}\cdots e^{ta_nX_n}] \\ &= E[e^{a_1tX_1}]E[e^{a_2tX_2}]\cdots E[e^{a_ntX_n}] \\ &= \prod_{i=1}^n M_{X_i}(a_i t). \end{aligned}$$

□

Cor 5.9. *If $X_1, \cdots, X_n$ follow an i.i.d. distribution with m.g.f. $M(t)$, then*

1. *the m.g.f. of $Y = \sum_{i=1}^n X_i$ is*

$$M_Y(t) = \prod_{i=1}^n M(t) = [M(t)]^n;$$

2. *the m.g.f. of $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i$ is*

$$M_{\overline{X}}(t) = \prod_{i=1}^n M\left(\frac{t}{n}\right) = \left[M\left(\frac{t}{n}\right)\right]^n.$$

The degrees of freedom in chi-square distribution may be interpreted by the following results.

Thm 5.10. *Let $X_1, \cdots, X_n$ be independent chi-square r.v.s with $r_1, \cdots, r_n$ degrees of freedom, respectively. Then the distribution of the r.v. $Y = X_1 + \cdots + X_n$ is $\chi^2(r_1 + \cdots + r_n)$.*

Proof. By Theorem 5.8, the m.g.f. of Y is

$$\begin{aligned} M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) = (1-2t)^{-r_1/2}(1-2t)^{-r_2/2}\cdots(1-2t)^{-r_n/2} \\ &= (1-2t)^{-\sum r_i/2}, \qquad \text{with } \ t < 1/2. \end{aligned}$$

which is the m.g.f. of $\chi^2(r_1 + r_2 + \cdots + r_n)$. □

Cor 5.11. *Let* $Z_1, \cdots, Z_n$ *have i.i.d.* $N(0,1)$*. Then*

$$W := Z_1^2 + \cdots + Z_n^2 \sim \chi^2(n).$$

Cor 5.12. *If* $X_1, \cdots, X_n$ *are independent and have normal distributions* $N(\mu_i, \sigma_i^2)$, $i = 1, \cdots, n$, *respectively, then*

$$W = \sum_{i=1}^{n} \frac{(X_i - \mu_i)^2}{\sigma_i^2} \quad \sim \quad \chi^2(n).$$

Homework
§*5.4* *1, 3, 5, 7, 9, 11*

Attachment: *Scanned textbook pages of Section 5-4*

5.5 Random Functions Associated with Normal Distributions

Many distributions are related to normal distributions.

Thm 5.13. *If $X_1, \cdots, X_n$ are n independent normal distribution with $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, \cdots, n$, then the linear function*

$$Y := \sum_{i=1}^{n} c_i X_i \quad \sim \quad N\left(\sum_{i=1}^{n} c_i \mu_i,\ \sum_{i=1}^{n} c_i^2 \sigma_i^2\right).$$

Cor 5.14. *If $X_1, \cdots, X_n$ are i.i.d. with distribution $N(\mu, \sigma^2)$, then the distribution of the sample mean*

$$\overline{X} := \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) \quad \sim \quad N\left(\mu, \frac{\sigma^2}{n}\right).$$

Ex *1-2, p.247-248 (scanned file)*

Thm 5.15. *Let $X_1, \cdots, X_n$ be observations of a random sample of size n from the normal distribution $N(\mu, \sigma^2)$. Then the sample mean*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

and the sample variance,

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

are independent and

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

Ex *3, p.250 (scanned file)*

Thm 5.16 (Student's t distribution). *Let* $T = \dfrac{Z}{\sqrt{U/r}}$, *where* $Z \sim N(0,1)$, $U \sim \chi^2(r)$, *and* Z *and* U *are independent. Then* T *has a* t *distribution of* r *degrees of freedom, with p.d.f.*

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

Denote $T \sim t(r)$ *if* T *has a* t *distribution with* r *degrees of freedom. The graph of p.d.f. of a* t *distribution is similar to that of a normal distribution, except that a* t *distribution has heavier tails than a normal distribution does. Let* $t_\alpha(r)$ *denote the right-tail probability of size* α, *See Fig 5.5-2(b), p252 (scanned file).*

Ex *4, p.252 (scanned file)*

Homework
§*5.5* *1, 3, 5, 7, 17*

Attachment: *Scanned textbook pages of Section 5-5*

5.6 The Central Limit Theorem

If $X_1, \cdots, X_n$ are i.i.d. with mean μ and variance σ^2, then the mean $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ has a mean μ and a variance σ^2/n. As n increases, the variance of $\overline{X}$ depends on n and goes to 0. To adjust this, let

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}.$$

Then for any n, W has mean 0 and variance 1.

Ex *If $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$, then $\overline{X} \sim N(\mu, \sigma^2/n)$. Thus $W \sim N(0,1)$ regardless of n.*

Thm 5.17 (Central Limit Theorem)**.** *If $\overline{X}$ is the mean of a random sample $X_1, \cdots, X_n$ of size n from a distribution with a finite mean μ and a finite positive variance σ^2, then the distribution of*

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$$

is $N(0,1)$ in the limit as $n \to \infty$.

Ex *1-5, p.256-257 (scanned file)*

Homework
§*5.6*   *1, 3, 5, 13, 15*

Attachment: *Scanned textbook pages of Section 5-6*

5.7 Approximations for Discrete Distributions

We will illustrate how the normal distribution can be used to approximate probabilities for certain discrete distributions.

Ex *Let* $X_1, \cdots, X_n$ *be a random sample from a Bernoulli distribution with mean* $\mu = p$ *and variance* $\sigma^2 = p(1-p)$*. Then* $Y = \sum_{i=1}^{n} X_i$ *is* $b(n, p)$*. By the central limit theorem.*

$$W = \frac{Y - np}{\sqrt{np(1-p)}} = \frac{\overline{X} - p}{\sqrt{p(1-p)/n}}$$

is $N(0,1)$ *in the limit as* $n \to \infty$*. So when* n *is large,* Y *is approximately* $N[np, np(1-p)]$*. The approximation usually requires that* $np \geq 5$ *and* $n(1-p) \geq 5$*.*

Ex *1-6, p.263-266 (scanned file)*

Homework
§*5.7* *1, 3, 5, 7, 11*

Attachment: *Scanned textbook pages of Section 5-7*