Probability and Statistics Notes

Chapter Five

Jesse Crawford

Department of Mathematics
Tarleton State University

Spring 2011

Outline

Change of Variables in the Bivariate Case

Theorem

- *Suppose* X_1 *and* X_2 *are random variables with joint p.d.f.* $f(x_1, x_2)$.
- *Let* $Y_1 = u_1(X_1, X_2)$ *and* $Y_2 = u_2(X_1, X_2)$.
- *Also, assume the transformation is 1-1 and satisfies certain regularity conditions analogous to those in section 5.1.*
- *Let* $X_1 = v_1(Y_1, Y_2)$ *and* $X_2 = v_2(Y_1, Y_2)$ *be the inverse mappings.*
- *Then the joint p.d.f. for* 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

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

The F Distribution

Definition

- Suppose U and V are independent, and
- $U \sim \chi^2(r_1)$ and $V \sim \chi^2(r_2)$.
- Then the random variable

  $$W = \frac{U/r_1}{V/r_2}$$

  is said to have an *F distribution* with r_1 and r_2 degrees of freedom, denoted $F(r_1, r_2)$.
- If $0 < \alpha < 1$, then $F_\alpha(r_1, r_2)$ is the critical value such that

  $$P[W \geq F_\alpha(r_1, r_2)] = \alpha.$$

Proposition

If $W \sim F(r_1, r_2)$, then the p.d.f. for W is

$$f(w) = \frac{(r_1/r_2)^{r_1/2}\Gamma[(r_1+r_2)/2]w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)[1+(r_1 w/r_2)]^{(r_1+r_2)/2}}.$$

Outline

Independence and Random Samples

Definition

- Suppose $X_1, \ldots, X_n$ are random variables with joint p.d.f. $f(x_1, \ldots, x_n)$, and
- let $f_i(x_i)$ be the p.d.f. of X_i, for $i = 1, \ldots, n$.
- Then $X_1, \ldots, X_n$ are *independent* if

$$f(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n).$$

- If these random variables all have the same distribution, they are said to be *identically distributed*.
- If $X_1, \ldots, X_n$ are independent and identically distributed (IID), then they are referred to as a *random sample of size n* from their common distribution.

Example

- A certain population of women have heights that are normally distributed,
- with mean 64 inches and standard deviation 2 inches.
- Let (X_1, X_2, X_3) be a random sample of size 3 from this population.
- Find the joint p.d.f. for (X_1, X_2, X_3).
- Find the probability that everyone's height in the sample exceeds 67 inches.

Proposition

- If $X_1, \ldots, X_n$ are independent, then for any sets $A_1, \ldots, A_n$,

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n)$$

- Also, for any functions $u_1, \ldots, u_n$,

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

Mean and Variance of a Linear Combination of R.V.'s

Theorem

- *Suppose* $X_1, \ldots, X_n$ *are independent R.V.'s*
- *with means* $\mu_1, \ldots, \mu_n$, *and*
- *variances* $\sigma_1^2, \ldots, \sigma_n^2$.
- *If* $a_1, \ldots, a_n \in \mathbb{R}$, *then*

$$E[a_1 X_1 + \cdots + a_n X_n] = a_1 \mu_1 + \cdots + a_n \mu_n, \text{ and}$$

$$Var[a_1 X_1 + \cdots + a_n X_n] = a_1^2 \sigma_1^2 + \cdots + a_n^2 \sigma_n^2.$$

Mean and Variance of the Sample Mean

Definition

- Let $X_1, \ldots, X_n$ be a random sample.
- The *sample mean* is

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

Proposition

- Let $X_1, \ldots, X_n$ be a random sample from a population with
- population mean μ and population variance σ^2.
- Then

$$E(\overline{X}) = \mu, \text{ and}$$

$$\text{Var}(\overline{X}) = \frac{\sigma^2}{n}.$$

Outline

m.g.f. of a Linear Combination

Theorem (5.4-1)

- *Suppose $X_1, \ldots, X_n$ are independent R.V.'s with*
- *moment-generating functions $M_{X_i}(t)$, for $i = 1, \ldots, n$.*
- *Then the moment-generating function of $Y = \sum_{i=1}^{n} a_i X_i$ is*

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

Example (5.4-2)

- Suppose $0 \leq p \leq 1$, and
- $X_1, \ldots, X_n$ all have a Bernoulli(p) distribution.
- Find the distribution of $Y = \sum_{i=1}^{n} X_i$.

Corollaries for Random Samples

Corollary (5.4-1)

- *Suppose* $X_1, \ldots, X_n$ *is a random sample*
- *from a distribution with m.g.f.* $M(t)$*.*
- *Then the m.g.f. of* $Y = \sum_{i=1}^{n} X_i$ *is*

$$M_Y(t) = [M(t)]^n, \textit{ and}$$

- *the m.g.f. of* $\overline{X}$ *is*

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

Example (5.4-3)

- Suppose (X_1, X_2, X_3) is a random sample from
- an exponential distribution with mean θ.
- Find the distributions of $Y = X_1 + X_2 + X_3$ and $\overline{X}$.

Theorem (5.4-2)

- *If* $X_1, \ldots, X_n$ *are independent, and*
- $X_i \sim \chi^2(r_i)$*, for each i, then*

$$X_1 + \cdots + X_n \sim \chi^2(r_1 + \cdots + r_n).$$

Corollary (5.4-2)

If $Z_1, \ldots, Z_n$ *are independent standard normal R.V.'s, then*

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

Corollary (5.4-3)

- *If* $X_1, \ldots, X_n$ *are independent, and each* $X_i \sim N(\mu_i, \sigma_i^2)$*, then*

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

Outline

Linear Combinations of Independent Normal R.V.'s

Theorem (5.5-1)

- *Suppose* $X_1, \ldots, X_n$ *are independent, and*
- $X_i \sim N(\mu_i, \sigma_i^2)$, *for* $i = 1, \ldots, n$.
- *Then*

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

Example (5.5-1)

- Suppose X_1 and X_2 are independent normal random variables,
- $X_1 \sim \mathrm{N}(693.2, 22820)$, and $X_2 \sim \mathrm{N}(631.7, 19205)$.
- Find $P(X_1 > X_2)$.

Sample Mean and Variance for a Normal Population

Theorem (5.5-2)

- *Let* $(X_1, \ldots, X_n)$ *be a random sample from* $N(\mu, \sigma^2)$*.*
- *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. Their distributions are*

$$\overline{X} \sim N(\mu, \sigma^2/n), \text{ and } S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1).$$

Example

- Consider a population of women whose heights are
- normally distributed with mean 64 inches
- and standard deviation 2 inches.
- For a sample of size $n = 10$, find $P(63 < \overline{X} < 65)$, and
- find constants a and b such that $P(a < S^2 < b) = 0.95$.
- Repeat the problem when $n = 81$.

Unbiasedness of $\overline{X}$ and S^2

- From Theorem 5.5-2, we have

$$E[\overline{X}] = \mu, \text{ and } E[S^2] = \sigma^2.$$

- $\overline{X}$, the sample mean, is used to *estimate* the population mean μ.
- S^2, the sample variance, is used to estimate the population variance σ^2.
- On average, each of these estimators are equal to the parameters they are intended to estimate.
- That is, $\overline{X}$ and S^2 are *unbiased*.

Remarks about Degrees of Freedom

- In the proof of Theorem 5.5-2, we noted that

$$\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n), \text{ and}$$

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

- Replacing the *parameter* μ with its *estimator* $\overline{X}$ resulted in a loss of one degree of freedom.
- There are many examples where the degrees of freedom is reduced *by one for each parameter being estimated*.

Student's t Distribution

Theorem (5.5-3)

- *Suppose Z and U are independent r.v.'s,*
- $Z \sim N(0,1)$*, and* $U \sim \chi^2(r)$.
- *Then,*
$$T = \frac{Z}{\sqrt{U/r}}$$
has a t distribution with r degrees of freedom, denoted $t(r)$.
- *The p.d.f. for a t distribution is*
$$f(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+t^2/r)^{(r+1)/2}}, \text{ for } t \in \mathbb{R}.$$

Relevance to Samples from Normal Distributions

Corollary

- *Suppose* $X_1, \ldots, X_n$ *is a random sample from* $N(\mu, \sigma^2)$*.*
- *Then*

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has a t distribution with $n - 1$ *degrees of freedom.*

Outline

The Central Limit Theorem

Theorem (5.6-1)

- *Suppose* $X_1, X_2, \ldots$ *is a sequence of IID random variables,*
- *from a distribution with finite mean* μ
- *and finite positive variance* σ^2*.*
- *Let* $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$*, for* $n = 1, 2, \ldots$
- *Then, as* $n \to \infty$*,*

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \Rightarrow N(0, 1).$$

Advanced texts:

- *Introduction to Mathematical Statistics*, 6th ed., by Hogg, McKean, and Craig.
- *Probability and Measure*, 3rd ed., by Billingsley.

Informal Statement of CLT

Informal CLT

- Suppose $X_1, \ldots, X_n$ is a random sample
- from a distribution with finite mean μ
- and finite positive variance σ^2.
- Then, if n is sufficiently large,

$$\overline{X} \approx \mathrm{N}(\mu, \sigma^2/n), \text{ and}$$

$$\sum_{i=1}^{n} X_i \approx \mathrm{N}(n\mu, n\sigma^2).$$

- Conventionally, values of $n \geq 30$ are usually considered sufficiently large, although this text applies the approximation for lower values of n, such as $n \geq 20$.

Examples

Example

- Consider a random sample of size 3000
- from a uniform distribution on the interval $[0, 1000]$.
- Find (approximately) $P(490 < \overline{X} < 510)$.
- Find
$$P\left(1,470,000 < \sum_{i=1}^{3000} X_i < 1,530,000\right).$$

Lottery Tickets

Example

Consider a \$1 scratch-off lottery ticket with the following prize structure:

Prize(\$)	Probability
0	0.80
2	0.15
10	0.05

- Find the expected profit/loss from buying a single ticket. Also find the standard deviation.
- What is the chance of breaking even if you buy one ticket?
- If you buy 100 tickets?
- If you buy 500 tickets?

Insurance

Example

- An auto insurance company has one million (statistically independent) customers.
- The annual costs incurred by an individual customer due to auto accidents are summarized below:

Cost($)	0	500	5,000	15,000
Probability	0.80	0.10	0.08	0.02

- Also, assume that each customer has at most one accident per year and has a $500 deductible.
- Find the expected value and variance of a single customer's claims.
- How much money must the company have to cover all of its customers' claims with 99% probability?

Outline

Normal Approximation to the Binomial Distribution

Proposition

- If $np \geq 5$ and $n(1-p) \geq 5$, then
- $b(n, p) \approx \mathrm{N}(np, np(1-p))$.

Example

- In a city with a population of 10 million people,
- 55% of the population approves of the mayor.
- In a random sample of size 2000,
- find the probability that the number of people who approve of the mayor is between 1060 and 1150 inclusive.

Continuity Correction

- When using this approximation to calculate probabilities,
- increase the width of the interval by 0.5 at each end.

Example

- Suppose $X \sim b(20, 0.3)$.
- Approximate the following probabilities:
  - $P(2 \leq X \leq 8)$
  - $P(2 < X < 8)$
  - $P(2 < X \leq 8)$

Normal Approximation to the Poisson Distribution

Proposition

- If n is sufficiently large,
- then $\text{Poiss}(n) \approx \text{N}(n, n)$.

Example

- A radioactive sample emits β-particles according to a Poisson process
- at an average rate of 35 per minute.
- Find the probability that the number of particles emitted
- in a 20 minute period exceeds 720.

Outline

- Know how to do all homework and quiz problems.
- Given the joint p.d.f. of two random variables, be able to determine
  - probabilities/expected values involving both random variables

$$P[(X, Y) \in A] = \int\int_A f(x, y)\, dydx, \text{ for any } A \subset \mathbb{R}^2.$$

$$E[u(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x, y)f(x, y)\, dydx.$$

  - marginal p.d.f.'s and probabilities/expected values involving only one of the variables

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)\, dy.$$

$$P(X \in A) = \int_A f_1(x)\, dx, \text{ for any } A \subseteq \mathbb{R}.$$

$$E[u(X)] = \int_{-\infty}^{\infty} u(x)f_1(x)\, dx.$$

- - conditional p.d.f.'s, conditional mean/variance, and conditional probabilities

$$g(x \mid y) = \frac{f(x, y)}{f_2(y)}.$$

$$E[u(X) \mid Y = y] = \int_{-\infty}^{\infty} u(x) g(x \mid y)\, dx.$$

$$\text{Var}(X \mid Y = y) = E(X^2 \mid Y = y) - E(X \mid Y = y)^2.$$

  - the covariance and correlation coefficient

$$\sigma_{XY} = \text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

  - the least squares regression line relating the variables

$$y = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

- Be able to find the expected value, variance, and probabilities involving Y, conditioning on $X = x$, when X and Y are jointly normal

$$\mu_{Y|x} = E(Y \mid X = x) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

$$\sigma^2_{Y|x} = \text{Var}(Y \mid X = x) = \sigma^2_Y(1 - \rho^2).$$

The conditional distribution of Y given $X = x$ is $\text{N}(\mu_{Y|x}, \sigma^2_{Y|x})$.

- Be able to find the distribution of $Y = u(X)$ using the distribution function technique or the change of variables formula.

$$G(y) = P(Y \leq y) = P(u(X) \leq y), \text{ and } g(y) = G'(y).$$

$$g(y) = f[v(y)]|v'(y)|, \text{ where } v = u^{-1}.$$

$$g(y_1, y_2) = f[v_1(y_1, y_2), v_2(y_1, y_2)] \left\| \begin{array}{cc} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{array} \right\|.$$

- Know the material in sections 5.3-5.6 and be able to solve related problems.