CS 660: Mathematical Foundations of Analytics

Dr. Francis Parisi

Pace University

Spring 2018

IMS Chapter 2 – Multivariate Distributions
FPA Sections 2.6, 2.7

Multivariate Distributions

(Joint) Distribution of Two Random Variables

- In the previous chapter we studied random variables taken one at a time
- We now mover on to the distribution of multiple random variables

Example

A coin is tossed three times and our interest is in the ordered number pair (number of H's on first two tosses, number of H's on all three tosses), where H and T represent, respectively, heads and tails. Let $\mathcal{C} = \{TTT, TTH, THT, HTT, THH, HTH, HHT, HHH\}$ denote the sample space. Let X_1 denote the number of H's on the first two tosses and X_2 denote the number of H's on all three flips. Then our interest can be represented by the pair of random variables (X_1, X_2).

Multivariate Distributions

Example (continued)

So for example, we toss the coin three times and get H, T, H. Then our ordered pair is $(X_1(HTH), X_2(HTH))$ which represents the outcome $(1,2)$, because we have one H on the first two tosses, and two H's on all three. The random variables X_1 and X_2 are real-valued functions from $\mathcal{C}$ to the sample space

$$\mathcal{D} = \{(0,0), (0,1), (1,1), (1,2), (2,2), (2,3)\}$$

Thus we have a vector function

$$(X_1, X_2) : \mathcal{C} \to \mathcal{D}$$

Multivariate Distributions

Definition (Random Vector)

Given a random experiment with a sample space $\mathcal{C}$, consider two random variables X_1 and X_2, which assign to each element c of C one and only one ordered pair of numbers $X_1(c) = x_1, X_2(c) = x_2$. Then we say that (X_1, X_2) is a **random vector.** The **space** of (X_1, X_2) is the set of ordered pairs $\mathcal{D} = \{(x_1, x_2) : x_1 = X_1(c), x_2 = X_2(c), c \in \mathcal{C}\}$.

Multivariate Distributions

- For some subset A of $\mathcal{D}$ we say A is an event
- We denote the probability of A as $P_{X_1,X_2}[A]$
- We define this probability in terms of the cumulative distribution function (cdf)

  $$F_{X_1,X_2}(x_1,x_2) = P\left[\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}\right]$$

  which we write as

  $$P\left[X_1 \leq x_1, X_2 \leq x_2\right]$$

- We call this the joint CDF of (X_1, X_2)
- Random vectors can be discrete or continuous

Multivariate Distributions

A random vector (X_1, X_2) is discrete if its space $\mathcal{D}$ is finite or countable

If discrete, then both X_1 and X_2 are discrete and the joint probability mass function (pmf) is

$$p_{X_1,X_2}(x_1, x_2) = P\left[X_1 = x_1, X_2 = x_2\right]$$

for all $(x_1, x_2) \in \mathcal{D}$

As with random variables, the pmf has the following properties

$$(i)\ 0 \leq p_{X_1,X_2}(x_1, x_2) \leq 1 \text{ and } (ii)\ \sum_{\mathcal{D}}\sum p_{X_1,X_2}(x_1, x_2) = 1$$

For any event $B \in \mathcal{D}$, we have

$$P\left[(X_1, X_2) \in B\right] = \sum_{B}\sum p_{X_1,X_2}(x_1, x_2)$$

Multivariate Distributions

Looking at the example before involving three coins, we can create a table to reflect the pmf of (X_1, X_2)

Support of X_2

Support of X_1	0	1	2	3
0	$\frac{1}{8}$	$\frac{1}{8}$	0	0
1	0	$\frac{2}{8}$	$\frac{2}{8}$	0
2	0	0	$\frac{1}{8}$	$\frac{1}{8}$

From the table we can easily find for example,

$$P\left[(X_1 \geq 2, X_2 \geq 2)\right] = p(2,2) + p(2,3) = 2/8$$

The support of a discrete random variable is the set of points in the space of X_1, X_2 such that $p(x_1, x_2) > 0$

Multivariate Distributions

For a continuous random vector (X_1, X_2) we represent the cdfs as integrals of non-negative functions
We can express $F_{X_1,X_2}(x_1, x_2)$ as

$$F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(w_1, w_2) dw_1 dw_2$$

for all $(x_1, x_2 \in \mathbb{R})$
The integrand is the **joint probability density function** of (X_1, X_2) which is

$$f_{X_1,X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1,X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$

Multivariate Distributions

Remark

When it is clear from the context we can drop the subscripts from the joint cdfs, pmfs, and pdfs. We may use notation like f_{12} instead of f_{X_1,X_2}. Lastly we will often use X, Y to represent a random vector besides (X_1, X_2)

Multivariate Distributions

Example (2.1.2)

Consider a continutous random vector X, Y which is uniformly distributed over the unit circle in $\mathbb{R}^2$. Since the are of the unit circle is π, the joint pdf is

$$f(x,y) = \begin{cases} \frac{1}{\pi} & -1 < y < 1, -\sqrt{1-y^2} < x < \sqrt{1-y^2} \\ 0 & \text{elsewhere.} \end{cases}$$

Suppose A is the interior of the circle with radius 1/2, then

$$P\left[(X,Y) \in A\right] = \frac{\pi\left(\frac{1}{2}\right)^2}{\pi} = \frac{1}{4}$$

Multivariate Distributions

Marginal Distributions

- Suppose (X_1, X_2) is a random vector
- Then both X_1 and X_2 are random variables
- Their distributions are called **marginal** distributions and we can derive the marginal distributions from the joint distribution
- The event that defines the cdf of X_1 at x_1 is $\{X_1 \leq x_1\}$
- Thus

$$\begin{aligned}\{X_1 \leq x_1\} &= \{X_1 \leq x_1\} \cap \{-\infty < X_2 < \infty\} \\ &= \{X_1 \leq x_1, -\infty < X_2 < \infty\}\end{aligned}$$

- That is, to find the probability that $X_1 \leq x_1$ we keep x_1 fixed and integrate (sum in the discrete case) over all x_2

Multivariate Distributions

		Support of X_2				
		0	1	2	3	$p_{X_1}(x_1)$
Support of X_1	0	$\frac{1}{8}$	$\frac{1}{8}$	0	0	$\frac{2}{8}$
	1	0	$\frac{2}{8}$	$\frac{2}{8}$	0	$\frac{4}{8}$
	2	0	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$
	$p_{X_2}(x_2)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

Multivariate Distributions

That is for a discrete random vector...

$$
\begin{aligned}
F_{X_1}(x_1) &= \sum_{w_1 \le x_1, -\infty < x_2 < \infty} \sum p_{X_1,X_2}(w_1, x_2) \\
&= \sum_{w_1 \le x_1} \left\{ \sum_{x_2 < \infty} p_{X_1,X_2}(w_1, x_2) \right\}
\end{aligned}
$$

and the quantity in the braces is the pmf of X_1

$$
p_{X_1}(x_1) = \sum_{x_2 < \infty} p_{X_1,X_2}(x_1, x_2)
$$

Multivariate Distributions

Similarly for a continuous random vector ...

$$
\begin{aligned}
F_{X_1}(x_1) &= \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} f_{X_1,X_2}(w_1, x_2) dx_2 dw_1 \\
&= \int_{-\infty}^{x_1} \left\{ \int_{-\infty}^{\infty} f_{X_1,X_2}(w_1, x_2) dx_2 \right\} dw_1
\end{aligned}
$$

and the quantity in the braces is the pdf of X_1

$$
f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) dx_2
$$

Multivariate Distributions

Expectations

- Suppose X_1, X_2 is a continuous random vector and $Y = g(X_1, X_2)$
- Then $\mathbb{E}[Y]$ exists if

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|g(x_1,x_2)|f_{X_1,X_2}(x_1,x_2)dx_1dx_2 < \infty$$

- Then

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x_1,x_2)f_{X_1,X_2}(x_1,x_2)dx_1dx_2$$

For the discrete case we have the analogous result

Multivariate Distributions

The Expectation Operator, $\mathbb{E}$ is a linear operator

For all real numbers k_1 and k_2 we have

$$\mathbb{E}[k_1 Y_1 + k_2 Y_2] = k_1\, \mathbb{E}[Y_1] + k_2\, \mathbb{E}[Y_2]$$

Multivariate Distributions

Definition (Expected Value of A Random Vector)

Let $X = (X_1, X_2)$ be a random vector. Then the expected value of X exists if the expectations of X_1 and X_2 exist. If it exists, then the expected value is given by

$$E[X] = \begin{bmatrix} E[X_1] \\ E[X_2] \end{bmatrix}$$

- In our discussion of univariate random variables we touched on the transformation of a random variable
- Not surprisingly this applies to bivariate random variable as well

Multivariate Distributions

Example (2.2.4)

Let $Y_1 = \frac{1}{2}(X_1 - X_2)$ where X_1 and X_2 have the joint pdf

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \frac{1}{4}\exp(-\frac{x_1+x_2}{2}) & 0 < x_1 < \infty,\ 0 < x_2 < \infty \\ 0 & \text{elsewhere} \end{cases}$$

Let $Y_2 = X_2$ so that $y_1 = \frac{1}{2}(x_1 - x_2)$, $y_2 = x_2$ or $x_1 = 2y_1 + y_2$, $x_2 = y_2$ define a one-to-one transformation from $\mathcal{S} = \{(x_1,x_2) : 0 < x_1 < \infty,\ 0 < x_2 < \infty\}$ onto $\mathcal{T} = \{(y_1,y_2) : -2y_1 < y_2,\ 0 < y_2 < \infty,\ -\infty < y_1 < \infty\}$.

Multivariate Distributions

Example (2.2.4 continued)

Recall in the univariate case we had dx/dy in the integral

For the multivariate case we need the **Jacobian** of the transformation, which for the bivariate case is

$$J = \begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} \end{vmatrix}$$

In this case

$$J = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = 2$$

Multivariate Distributions

Example (2.2.4 Continued)

The joint pdf of Y_1 and Y_2 is

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \frac{|2|}{4} e^{-y_1-y_2} & (y_1,y_2) \in \mathcal{T} \\ 0 & \text{elsewhere} \end{cases}$$

Finally we get the pdf for Y_1

$$f_{Y_1}(y_1) = \begin{cases} \int_{-2y_1}^{\infty} \frac{1}{2} e^{-y_1-y_2} dy_2 = \frac{1}{2} e^{y_1} & -\infty < y_1 < 0 \\ \\ \int_{0}^{\infty} \frac{1}{2} e^{-y_1-y_2} dy_2 = \frac{1}{2} e^{-y_1} & 0 \le y_1 < \infty \end{cases}$$

or

$$f_{Y_1}(y_1) = \frac{1}{2} e^{-|y_1|},\ -\infty < y_1 < \infty$$

This is the Laplace distribution aka the **double exponential** pdf

For X_1 and X_2 discrete random variables with joint pmf $p_{X_1,X_2}(x_1,x_2)$, with marginal pmfs $p_{X_1}(x_1)$, and $p_{X_2}(x_2)$

Then the **conditional pmf** of X_2 is

$$p_{X_2|X_1}(x_2|x_1) = \frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_1}(x_1)}$$

and the conditional pmf of X_1 is

$$p_{X_1|X_2}(x_1|x_2) = \frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_2}(x_2)}$$

Multivariate Distributions

We have analogous results for continuous random variables

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}$$

and

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}$$

Let's simplify the notation so we have $f_{1|2}(x_1|x_2),\ f_{2|1}(x_2|x_1),$ $f_1(x_1),$ and $f_2(x_2)$

Multivariate Distributions

If we wish to find the conditional probability that $a < X_2 < b$, given $X_1 = x_1$ then

$$P\left[a < X_2 < b | X_1 = x_1\right] = \int_a^b f_{2|1}(x_2|x_1)dx_2$$

likewise

$$P\left[c < X_1 < d | X_2 = x_2\right] = \int_c^d f_{1|2}(x_1|x_2)dx_1$$

Multivariate Distributions

Using the concept of conditional distribution we define **conditional expectation** and **conditional variance**

- If $u(X_2)$ is a function of X_2 the **conditional expectation** of $u(X_2)$ given $X_1 = x_1$ is

$$\mathbb{E}[u(X_2)|x_1] = \int_{-\infty}^{\infty} u(x_2) f_{2|1}(x_2|x_1) dx_2$$

- The **conditional variance** is

$$\mathbb{E}\{[X_2 - \mathbb{E}(X_2|x_1)]^2|x_1\}$$

or simply

$$\text{Var}\,(X_2|x_1) = \mathbb{E}[X_2^2|x_1] - (\mathbb{E}[X_2|x_1])^2$$

Multivariate Distributions

Example (2.3.1)

Let X_1 and X_2 have joint pdf

$$f(x_1, x_2) = \begin{cases} 2 & 0 < x_1 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

The marginal pdfs are

$$f_1(x_1) = \begin{cases} \int_{x_1}^{1} 2dx_2 = 2(1 - x_1) & 0 < x_1 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$f_2(x_2) = \begin{cases} \int_{0}^{x_2} 2dx_1 = 2x_2 & 0 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Multivariate Distributions

Example (2.3.1 continued)

The conditional pdf of X_1 given $X_2 = x_2$, $0 < x_2 < 1$ is

$$f_{1|2}(x_1|x_2) = \begin{cases} \frac{2}{2x_2} = \frac{1}{x_2} & 0 < x_1 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the conditional mean and conditional variance of $X_1 | X_2 = x_2$.

Multivariate Distributions

Solution

$$\begin{aligned}
\mathbb{E}(X_1|x_2) &= \int_{-\infty}^{\infty} x_1 f_{1|2}(x_1|x_2)dx_1 \\
&= \int_0^{x_2} x_1 \left(\frac{1}{x_2}\right) dx_1 \\
&= \frac{x_2}{2},\ 0 < x_2 < 1
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}\left(X_1|x_2\right) &= \int_0^{x_2} \left(x_1 - \frac{x_2}{2}\right)^2 \left(\frac{1}{x_2}\right) dx_1 \\
&= \frac{x_2^2}{12},\ 0 < x_2 < 1
\end{aligned}$$

Multivariate Distributions

Let's see the effect of conditioning on finding the probability that $0 < X_1 < \frac{1}{2}$

$$P\left[0 < X_1 < \tfrac{1}{2}\middle| X_2 = \tfrac{3}{4}\right] = \int_0^{1/2} f_{1|2}\left(x_1 \middle| \tfrac{3}{4}\right) dx_1 = \int_0^{1/2} \left(\tfrac{4}{3}\right) dx_1 = \frac{2}{3}$$

however,

$$P\left[0 < X_1 < \tfrac{1}{2}\right] = \int_0^{1/2} f_1(x_1) dx_1 = \int_0^{1/2} 2(1 - x_1) dx_1 = \frac{3}{4}$$

So $P\left[0 < X_1 < \tfrac{1}{2}\middle| X_2 = \tfrac{3}{4}\right] \neq P\left[0 < X_1 < \tfrac{1}{2}\right]$

Definition

Let the random variables X_1 and X_2 have the joint pdf $f(x_1, x_2)$ [joint pmf $p(x_1, x_2)$] and the marginal pdfs [pmfs] $f_1(x_1)$ [$p_1(x_1)$] and $f_2(x_2)$ [$p_2(x_2)$], respectively. The random variables X_1 and X_2 are said to be independent if, and only if, $f(x_1, x_2) \equiv f_1(x_1)f_2(x_2)$ [$p(x1, x2) \equiv p1(x1)p2(x2)$]. Random variables that are not independent are said to be dependent.

Multivariate Distributions

Theorem

Let (X_1, X_2) have the joint cdf $F(x_1, x_2)$ and let X_1 and X_2 have the marginal cdfs $F_1(x_1)$ and $F_2(x_2)$, respectively. Then X_1 and X_2 are independent if and only if $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ for all $(x_1, x_2) \in R^2$.

Theorem

The random variablesX_1 and X_2 are independent random variables if and only if the following condition holds,

$$P\left[a < X_1 \leq b, c < X_2 \leq d\right] = P\left[a < X_1 \leq b\right] P\left[c < X_2 \leq d\right]$$

for every $a < b$ and $c < d$, where a, b, c, and d are constants.

Multivariate Distributions

Theorem

Suppose X_1 and X_2 are independent and that $\mathbb{E}(u(X_1))$ and $\mathbb{E}(v(X_2))$ exist. Then

$$\mathbb{E}[u(X_1)v(X_2)] = \mathbb{E}[u(X_1)]\,\mathbb{E}[v(X2)].$$

What's the point? For two independent random variables

- The joint CDF equals the product of the two marginal CDFS
- The joint probability equals the product of the two individual probabilities
- The expected value of the product if the product of the expected values

Multivariate Distributions

Exercise

Show that the random variables X_1 and X_2 with joint pdf

$$f(x_1, x_2) = \begin{cases} 12x_1x_2(1-x_2) & 0 < x_1 < 1,\ 0 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

are independent.

Multivariate Distributions

Solution

Multivariate Distributions

Correlation Coefficients

- If two random variables are independent then we discussed some properties that result
- hat if they are dependent? How do we measure ther degree of association between the two?
- We'll talk about the **covariance** of X, Y and then the **correlation** between the two

Multivariate Distributions

Definition (Covariance)

Let (X, Y) have a joint distribution. Denote the means of X and Y respectively by μ_1 and μ_2 and their variances by σ_1^2 and σ_2^2. The **covariance** of (X, Y) is denoted $\mathrm{Cov}(X, Y)$ and is defined as

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mu_1)(Y - \mu_2)]. \tag{2.1}$$

Because $\mathbb{E}$ is a linear operator we get

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mu_1\mu_2$$

which is easier to work with

Multivariate Distributions

Correlation is a standardized version of covariance and is more interpretable...

Definition

If each of σ_1 and σ_2 is positive then the **correlation coefficient** between X and Y is

$$\rho = \frac{\mathbb{E}[(X - \mu_1)(Y - \mu_2)]}{\sigma_1 \sigma_2} = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2}$$

Note from (2.1) that

$$\mathbb{E}[XY] = \mu_1 \mu_2 + \text{Cov}(X, Y) = \mu_1 \mu_2 + \rho \sigma_1 \sigma_2$$

Example (2.5.2)

Let the random variables X and Y have the joint pdf

$$f(x,y) = \begin{cases} x+y & 0<x<1,\ 0<y<1 \\ 0 & \text{elsewhere.} \end{cases}$$

Find ρ.

Multivariate Distributions

Solution

$$\mu_1 = \mathbb{E}[X] = \int_0^1 \int_0^1 x(x+y)dxdy = \frac{7}{12}$$

and

$$\sigma_1^2 = \mathbb{E}[X^2] - \mu_1^2 = \int_0^1 \int_0^1 x^2(x+y)dxdy - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

Similarly, $\mu_2 = \mathbb{E}[Y] = \frac{7}{12}$ and $\sigma_2^2 = \mathbb{E}[Y^2] - \mu_2^2 = \frac{11}{144}$
The covariance is

$$\mathbb{E}[XY] - \mu_1\mu_2 = \int_0^1 \int_0^1 xy(x+y)dxdy - \left(\frac{7}{12}\right)^2 = -\frac{1}{144}$$

Finally, $$\rho = \frac{-\frac{1}{144}}{\sqrt{(\frac{11}{144})(\frac{11}{144})}} = -\frac{1}{11}$$

Multivariate Distributions

Theorem (2.5.2)

If X and Y are independent random variables the $\mathrm{Cov}(X, Y) = 0$ and therefore, $\rho = 0$.

Multivariate Distributions

Theorem (2.5.3)

Suppose (X, Y) have a joint distribution with the variances of X and Y finite and positive. Denote the means and variances of X and Y by μ_1, μ_2 and σ_1^2, σ_2^2 respectively, and let ρ be the correlation coefficient between X and Y. If $\mathbb{E}[Y|X]$ is linear in X then

$$\mathbb{E}[Y|X] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X - \mu_1)$$

and

$$\mathbb{E}[\mathrm{Var}\,(Y|X)] = \sigma_2^2(1 - \rho^2).$$

We will now summarize some results on linear combinations of random variables

Let $(X_1, X_2, \ldots, X_n)'$ denote a random vector. Then we can write linear combinations of these random variables as

$$T = \sum_{i=1}^{n} a_i X_i,$$

for constants $a_i, \ldots, a_n$.

Multivariate Distributions

Theorem (2.8.1)

Suppose $T = \sum_{i=1}^{n} a_i X_i$, *as before and* $\mathbb{E}(X_i) = \mu_i$ *for* $i = 1, \ldots, n$.
Then

$$\mathbb{E}(T) = \sum_{i=1}^{n} a_i \mu_i$$

Multivariate Distributions

Theorem (2.8.2)

Suppose $T = \sum_{i=1}^{n} a_i X_i$, *and* $W = \sum_{i=1}^{m} b_i Y_i$, *for random variables* Y_i *and constants* b_i. *If* $\mathbb{E}[X_i^2] < \infty$ *and* $\mathbb{E}[Y_j^2] < \infty$ *for* $i = 1, \ldots, n$, *and* $j = 1, \ldots, m$, *then*

$$\operatorname{Cov}(T, W) = \sum_{i=1}^{n} \sum_{i=1}^{m} a_i b_j \operatorname{Cov}(X_i, Y_j).$$

Multivariate Distributions

Corollary (2.8.1)

Let $T = \sum_{i=1}^{n} a_i X_i$. *Provided* $\mathbb{E}[X_i^2] < \infty$ *for* $i = 1, \ldots, n$,

$$\mathrm{Var}\,(T) = \mathrm{Cov}(T, T) = \sum_{i=1}^{n} a_i^2 \mathrm{Var}\,(X_i) + 2 \sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

Multivariate Distributions

Corollary (2.8.2)

If $X_i, \ldots, X_n$ *are independent random variables and* $\text{Var}(X_i) = \sigma_i^2$, *for* $i = 1, \ldots, n$, *then*

$$\text{Var}(T) = \sum_{i=1}^{n} a_i^2 \sigma_i^2$$

Multivariate Distributions

Definition (2.8.1)

If the random variables $X_1, \ldots, X_n$ are independent and identically distributed (*iid*) i.e., each X_i has the same distribution, then we say that these random variables constitute a **random sample** of size n from that common distribution.

Multivariate Distributions

Example (2.8.1 – Sample Mean)

Let $X_1, \ldots, X_n$ be *iid* random variables with common mean μ and variance σ^2. The **sample mean** is defined by $\overline{X} = n^{-1} \sum_{i=1}^{n} X_i$. This is a linear combination of the sample observations with $a_i \equiv n^1$; hence, by Theorem 2.8.1 and Corollary 2.8.2, we have $\mathbb{E}(\overline{X}) = \mu$ and $\mathrm{Var}\left(\overline{X}\right) = \frac{\sigma^2}{n}$.

Because $\mathbb{E}(\overline{X}) = \mu$, we say $\overline{X}$ is an **unbiased estimator** of μ.

Multivariate Distributions

Definition (2.8.2 – Sample Variance)

Define the **sample variance** by

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

the second equality follows after some algebra. Using the above theorems, the results of the last example, and the fact that $\mathbb{E}(X^2) = \sigma^2 + \mu^2$, and $\mathbb{E}[\overline{X}^2] = (\sigma^2/n) + \mu^2$ we have the following:

$$\begin{aligned}
\mathbb{E}(S^2) &= (n-1)^{-1} \left(\sum_{i=1}^{n} \mathbb{E}[X_i^2] - n\,\mathbb{E}[\overline{X}^2] \right) \\
&= (n-1)^{-1} \{ n\sigma^2 + n\mu^2 - n[(\sigma^2/n) + \mu^2] \} \\
&= \sigma^2.
\end{aligned}$$

Hence, S^2 is an unbiased estimator of σ^2.

References

Hogg, R. V., McKean, J. W. and Craig, A. T. (2018).
Introduction to Mathematical Statistics.
Pearson, Boston, eighth edition.

Ross, S. (2010).
A First Course in Probability.
Prentice Hall, Upper Saddle River, eigth edition.

Wu, J. and Coggeshall, S. (2012).
Foundations of Predictive Analytics.
CRC Press, Boca Raton, FL.