Chapter 4

Multiple Random Variables

Joint and Marginal Distributions

Outline

Multivariate Distributions

We will now consider more than one random variable at a time. As we shall see, developing the theory of multivariate distributions will allow us to consider situations that model the actual collection of data and form the foundation of inference based on those data.

Discrete Random Variables

As with univariate random variables, we compute probabilities by adding the appropriate entries in the table.

$$P\{(X_1, X_2) \in B\} = \sum_{(x_1,x_2)\in B} f_{X_1,X_2}(x_1, x_2).$$

As before, the mass function has two basic properties.

- $f_{X_1,X_2}(x_1, x_2) \geq 0$ for all x_1 and x_2.
- $\sum_{x_1,x_2} f_{X_1,X_2}(x_1, x_2) = 1$.

The distribution of an individual random variable is call the marginal distribution. The marginal mass function for X_1 is found by summing over the appropriate column and the marginal mass function for X_2 can be found be summing over the appropriate row.

Discrete Random Variables

Example. For X_1 and X_2 each having finite range, we can display the mass function in a table.

		x_1				
		0	1	2	3	4
	0	0.02	0.02	0	0.10	0
	1	0.02	0.04	0.10	0	0
x_2	2	0.02	0.06	0	0.10	0
	3	0.02	0.08	0.10	0	0.05
	4	0.02	0.10	0	0.10	0.05

Exercise. Find

1. $P\{X_1 = X_2\}$. 0.11
2. $P\{X_1 + X_2 \leq 3\}$. 0.40
3. $P\{X_1 X_2 = 0\}$. 0.22
4. $P\{X_1 = 3\}$. 0.30

Discrete Random Variables

For the marginal mass functions;

		x_1					
		0	1	2	3	4	$f_{X_2}(x_2)$
x_2	0	0.02	0.02	0	0.10	0	0.14
	1	0.02	0.04	0.10	0	0	0.16
	2	0.02	0.06	0	0.10	0	0.18
	3	0.02	0.08	0.10	0	0.05	0.25
	4	0.02	0.10	0	0.10	0.05	0.27
	$f_{X_1}(x_1)$	0.10	0.30	0.20	0.30	0.10	

The definition of expectation in the case of a finite sample space S is a straightforward generalization of the univariate case.

$$Eg(X_1, X_2) = \sum_{x_1, x_2} g(x_1, x_2) f_{X_1, X_2}(x_1, x_2).$$

Exercise Compute EX_1X_2 in the example above.

Continuous Random Variables

For continuous random variables, we have the notion of the joint (probability) density function

$$f_{X_1,X_2}(x_1,x_2)\Delta x_1 \Delta x_2 \approx P\{x_1 < X_1 \le x_1 + \Delta x_1, x_2 < X_2 \le x_2 + \Delta x_2\}.$$

We can write this in integral form as

$$P\{(X_1,X_2) \in B\} = \int\int_B f_{X_1,X_2}(x_1,x_2)\, dx_2 dx_1.$$

The basic properties of the joint density function are

- $f_{X_1,X_2}(x_1,x_2) \ge 0$ for all x_1 and x_2.
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x,x_2)\, dx_2 dx_1 = 1.$

Continuous Random Variables

The joint cumulative distribution function is defined as

$$F_{X_1,X_2}(x_1, x_2) = P\{X_1 \le x_1, X_2 \le x_2\}.$$

For the case of continuous random variables, we have

$$F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1,X_2}(s_1, s_2)\, ds_2 \; ds_1.$$

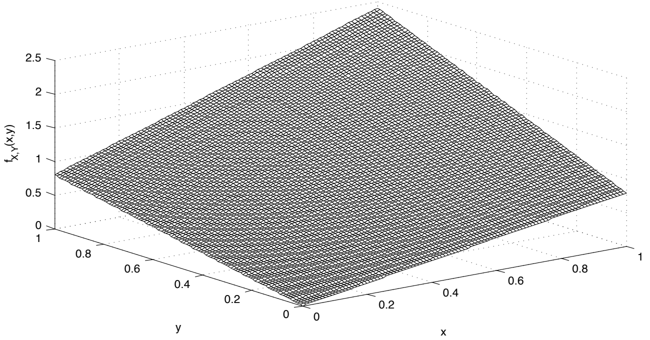
By two applications of the fundamental theorem of calculus, we find that

$$\frac{\partial}{\partial x_1} F_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{x_2} f_{X_1,X_2}(x_1, s_2)\, ds_2 \quad \text{and} \quad \frac{\partial^2}{\partial x_1 \partial x_2} F_{X_1,X_2}(x_1, x_2) = f_{X_1,X_2}(x_1, x_2).$$

Continuous Random Variables

Example. Let (X_1, X_2) have joint density

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} c(x_1x_2 + x_1 + x_2) & \text{for } 0 \le x_1 \le 1, 0 \le x_2 \le 1, \\ 0 & \text{otherwise.} \end{cases}$$



Continuous Random Variables

Then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\,dx_2dx_1 = \int_0^1\int_0^1 c(x_1x_2+x_1+x_2)\,dx_2dx_1$$

$$= c\int_0^1 \left(\frac{1}{2}x_1x_2^2 + x_1x_2 + \frac{1}{2}x_2^2\right)\Big|_0^1 dx_1 = c\int_0^1\left(\frac{3}{2}x_1+\frac{1}{2}\right)\,dx_1$$

$$= c\left(\frac{3}{4}x_1^2+\frac{1}{2}x_1\right)\Big|_0^1 = \frac{5c}{4}$$

and $c = 4/5$

$$P\{X_1 \geq X_2\} = \int_0^1\int_0^{x_1}\frac{4}{5}(x_1x_2+x_1+x_2)\,dx_2dx_1 = \frac{4}{5}\int_0^1\left(\frac{1}{2}x_1x_2^2+x_1x_2+\frac{1}{2}x_2^2\right)\Big|_0^{x_1} dx_1$$

$$= \frac{4}{5}\int_0^1\left(\frac{1}{2}x_1^3+\frac{3}{2}x_1^2\right)\,dx_1 = \frac{4}{5}\left(\frac{1}{8}x_1^4+\frac{1}{2}x_1^3\right)\Big|_0^1 = \frac{4}{5}\cdot\frac{5}{8} = \frac{1}{2}.$$

Continuous Random Variables

The formula for expectations can be obtained from discrete random variables using a similar limiting argument to obtain a Riemann sum for a multivariate definite integral. Thus,

$$Eg(X_1, X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1, x_2) f_{X_1,X_2}(x_1, x_2)\ dx_2 dx_1.$$

Exercise. For the density above, find EX_1X_2.

$$\begin{aligned}
EX_1X_2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f_{X_1,X_2}(x_1, x_2)\,dx_2dx_1 = \frac{4}{5}\int_0^1\int_0^1 x_1x_2(x_1x_2 + x_1 + x_2)\,dx_2dx_1 \\
&= \frac{4}{5}\int_0^1\int_0^1 (x_1^2x_2^2 + x_1^2x_2 + x_1x_2^2)\,dx_2dx_1 = \frac{4}{5}\int_0^1 \left(\frac{1}{3}x_1^2x_2^3 + \frac{1}{2}x_1^2x_2^2 + \frac{1}{3}x_1x_2^3\right)\Big|_0^1 dx_1 \\
&= \frac{4}{5}\int_0^1 \left(\frac{5}{6}x_1^2 + \frac{1}{3}x_1\right)\,dx_1 = \frac{4}{5}\left(\frac{5}{18}x_1^3 + \frac{1}{6}x_1^2\right)\Big|_0^1 = \frac{4}{5}\frac{5+3}{18} = \frac{16}{45}
\end{aligned}$$

Independent Random Variables

We say that two random variables X_1 and X_2 are independent if for any (measurable) sets B_1 and B_2, the events $\{X_1 \in B_1\}$ and $\{X_2 \in B_2\}$ are independent.

$$P\{X_1 \in B_1, X_2 \in B_2\} = P\{X_1 \in B_1\}P\{X_2 \in B_2\}.$$

For the particular choice of $B_1 = (-\infty, x_1]$ and $B_2 = (-\infty, x_2]$, we have that

$$F_{X_1,X_2}(x_1, x_2) = P\{X_1 \leq x_1, X_2 \leq x_2\} = P\{X_1 \leq x_1\}P\{X_2 \leq x_2\} = F_{X_1}(x_1)F_{X_2}(x_2).$$

In words, the joint cumulative probability distribution function is the product of the marginal distribution functions.

Independent Random Variables

For continuous random variables, we take partial derivatives to find that

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

and the joint density function is the product of the marginal density functions.

Similarly, for discrete random variables, take $B_1 = \{x_1\}$ and $B_2 = \{x_2\}$ to obtain

$$f_{X_1,X_2}(x_1,x_2) = P\{X_1 = x, X_2 = x_2\} = P\{X_1 = x_1\} P\{X_2 = x_2\} = f_{X_1}(x_1) f_{X_2}(x_2)$$

and the joint mass function is the product of the marginal mass functions.

Expectation

For both continuous and discrete random variables, we can write the expectation as a double integral

$$Eg(X_1, X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1, x_2)\, dF_{X_1,X_2}(x_1, x_2).$$

If $g(x_1, x_2) = g_1(x_1)g_2(x_2)$ and X_1 and X_2 are independent,then the identity above becomes

$$\begin{aligned} Eg_1(X_1)g_2(X_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x_1)g_2(x_2)\, dF_{X_1}(x_1)dF_{X_2}(x_2) \\ &= \left(\int_{-\infty}^{\infty} g_1(x_1)\, dF_{X_1}(x_1)\right)\left(\int_{-\infty}^{\infty} g_2(x_2)\, dF_{X_2}(x_2)\right) = Eg_1(X_1)Eg_2(X_2). \end{aligned}$$

and the expectation of the product and the expectation of the product is equal to the product of the expectations.

Variance

For independent X_1 and X_2

$$
\begin{aligned}
\mathrm{Var}(X_1 + X_2) &= E[((X_1 + X_2) - (\mu_{X_1} + \mu_{X_2}))^2] = E[((X_1 - \mu_{X_1}) + (X_2 - \mu_{X_2}))^2] \\
&= E[(X_1 - \mu_{X_1})^2] + 2E[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})] + E[(X_2 - \mu_{X_2})^2] \\
&= \mathrm{Var}(X_1) + 2E[X_1 - \mu_{X_1}]E[X_2 - \mu_{X_2}] + \mathrm{Var}(X_2) \\
&= \mathrm{Var}(X_1) + 0 + \mathrm{Var}(X_2)
\end{aligned}
$$

and the variance of the sum is the sum of the variances.

Exercise. For independent $X_i, 1 \leq i \leq n$,

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)$$

Example

For independent $X_i, 1 \le i \le n$, values on throws of a fair die. Then

$$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E[X_i] = \left(\frac{7}{2}\right)^n.$$

In addition,

$$\mathrm{Var}(X_i) = \frac{6^2 - 1}{12} = \frac{35}{12}.$$

Thus,

$$\mathrm{Var}(X_1 + \cdots + X_n) = \frac{35n}{12}.$$

Example

Exercise. For $X_i, 1 \le i \le n$, define the sample mean

$$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$$

- Then, if the X_i each have mean μ, then $E[\bar{X}] = \mu$,

$$E[\bar{X}] = E\left[\frac{1}{n}(X_1 + \cdots + X_n)\right] = \frac{1}{n}(EX_1 + \cdots + EX_n) = \frac{1}{n}n\mu = \mu$$

- Then, if the variables are independent each with variance σ^2, $\mathrm{Var}(\bar{X}) = \sigma^2/n$.

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{n}(X_1 + \cdots + X_n)\right) = \frac{1}{n^2}(\mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

Thus for n rolls of a fair die $E[\bar{X}] = 7/2$ and $\mathrm{Var}(\bar{X}) = 35/(12n)$.

Skewness

Now, let the $X_i, 1 \le i \le n$, be independent with a common distribution having mean μ, variance σ^2, and skewness

$$\gamma_1 = E\left[\left(\frac{X_i - \mu}{\sigma}\right)^3\right]$$

To find the skewness of $\bar{X}$, we standardize

$$\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^3 = \left(\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}\right)^3 = \frac{1}{n^{3/2}}\left(\sum_{i=1}^n \frac{X_i - \mu}{\sigma}\right)^3 = \frac{1}{n^{3/2}}\left(\sum_{i=1}^n X_i^*\right)^3$$

where X_i^* is the stardardization of X_i,

Skewness

Note that $EX_i^* = 0$, $E(X_i^*)^2 = 1$, and $E(X_i^*)^3 = \gamma_1$. Next, expand the cube of the sum on the X_i^*, take expectation and use its linearity. Then we find terms

- where the indices i, j, k all differ. Then,

$$E[X_i^* X_j^* X_k^*] = E[X_i^*]E[X_j^*]E[X_k^*] = 0 \cdot 0 \cdot 0 = 0.$$

- where exactly two indices are the same $i = j \neq k$. Then,

$$E[X_i^* X_j^* X_k^*] = E[(X_i^*)^2]E[X_k^*] = 1 \cdot 0 = 0.$$

- where the three indices are the same $i = j = k$. Then,

$$E[X_i^* X_j^* X_k^*] = E[(X_i^*)^3] = \gamma_1.$$

Thus, $E(\sum_{i=1}^n X_i^*)^3 = \sum_{i=1}^n E(X_i^*)^3 = n\gamma_1$, and the skewness of $\bar{X}$ is $\gamma_1/\sqrt{n}$