

ISEG - Lisbon School of Economics and Management
Statistics I

(Date of this version: 22/11/2017)

Handout 3 (Part 2) – Expected values and parameters

Issues covered:

- 3.12 Expected values of functions of random variables
- 3.13 Marginal expected values
- 3.14 Product of moments around the origin
- 3.15 Product moments about the mean
- 3.16 The covariance
- 3.17 The correlation coefficient
- 3.18 Moments of linear functions of random variables
- 3.19 Conditional expectations
- 3.20 The law of iterated expectations

3.12 Expected values of functions of random variables

Let $D_{(X,Y)}$ be the set of points of discontinuity of the joint cumulative distribution function $F_{X,Y}(x, y)$.

Definition: (Expected value of a function of a discrete two-dimensional random variable): Let (X, Y) be a discrete two-dimensional random variable with joint probability function $f_{X,Y}(x, y)$ and let $g(X, Y)$ be a function of (X, Y) , then

$$E[g(X, Y)] = \sum_{(x,y) \in D_{(X,Y)}} g(x, y) f_{X,Y}(x, y)$$

provided that $\sum_{(x,y) \in D_{(X,Y)}} |g(x, y)| f_{X,Y}(x, y) < +\infty$.

Remark: In the case $D_{(X,Y)}$ is countably infinite $D_{(X,Y)} = \{(x_i, y_j) : i = 1, 2, \dots; j = 1, 2, \dots\}$ we have

$$E[g(X, Y)] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} g(x_i, y_j) f_{X,Y}(x_i, y_j)$$

Definition: (Expected value of a function of a continuous two-dimensional random variable) Let (X, Y) be a continuous two-dimensional random variable with joint probability density function $f_{X,Y}(x, y)$ and let $g(X, Y)$ be a function of (X, Y) , then

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{X,Y}(x, y) dx dy$$

provided that $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |g(x, y)| f(x, y) dx dy < +\infty$.

3.13 Marginal expected values

Theorem: Let (X, Y) be a discrete two-dimensional random variable with joint probability function $f_{X,Y}(x, y)$:

1. If $g(X, Y) = h(X)$ that is $g(X, Y)$ only depends on X , then

$$\begin{aligned} E(g(X, Y)) &= E[h(X)] \\ &= \sum_{(x,y) \in D_{(X,Y)}} h(x) f_{X,Y}(x, y) = \sum_{x \in D_X} h(x) \sum_{y \in D_Y} f_{X,Y}(x, y) = \sum_{x \in D_X} h(x) f_X(x) \end{aligned}$$

provided that $\sum_{(x,y) \in D_{(X,Y)}} |h(x)| f_{X,Y}(x, y) < +\infty$.

2. If $g(X, Y) = v(Y)$ that is $g(X, Y)$ only depends on Y , then

$$E[v(Y)] = \sum_{(x,y) \in D_{(X,Y)}} v(y) f_{X,Y}(x, y) = \sum_{y \in D_Y} v(y) \sum_{x \in D_X} f_{X,Y}(x, y) = \sum_{y \in D_Y} v(y) f_Y(y)$$

provided that $\sum_{(x,y) \in D_{(X,Y)}} |v(y)| f_{X,Y}(x, y) < +\infty$.

Theorem: Let (X, Y) be a continuous two-dimensional random variable with joint probability function $f_{X,Y}(x, y)$:

1. If $g(X, Y) = h(X)$ that is $g(X, Y)$ only depends on X , then

$$E[h(X)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x) f_{X,Y}(x, y) dx dy = \int_{-\infty}^{+\infty} h(x) \left(\int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy \right) dx = \int_{-\infty}^{+\infty} h(x) f_X(x) dx$$

provided that $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |h(x)| f_{X,Y}(x, y) dx dy < +\infty$.

2. If $g(X, Y) = v(Y)$ that is $g(X, Y)$ only depends on Y , then

$$E[v(Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v(y) f_{X,Y}(x, y) dx dy = \int_{-\infty}^{+\infty} v(y) \left(\int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx \right) dy = \int_{-\infty}^{+\infty} v(y) f_Y(y) dy$$

provided that $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |v(y)| f_{X,Y}(x, y) dx dy < +\infty$.

Properties:

- $E[h(X) + v(Y)] = E[h(X)] + E[v(Y)]$ provided that $E[|h(X)|] < +\infty$, $E[|v(Y)|] < +\infty$
- $E\left[\sum_{i=1}^N X_i\right] = \sum_{i=1}^N E[X_i]$, where N is a finite integer, provided that $E[|X_i|] < +\infty$ for $i = 1, 2, \dots, N$.

3.14 Product of moments around the origin

Definition: The r th and s th product moment about the origin of the discrete random variables X and Y , denoted by $\mu'_{r,s}$ is the expected value of $X^r Y^s$, for $r = 1, 2, \dots$; $s = 1, 2, \dots$ which is given by

$$\mu'_{r,s} = E[X^r Y^s] = \sum_{(x,y) \in D_{(X,Y)}} x^r y^s f_{X,Y}(x, y)$$

in the case of discrete random variables.

Remark: In the case $D_{(X,Y)}$ is countably infinite we have

$$\mu'_{r,s} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i^r y_j^s f_{X,Y}(x_i, y_j)$$

Definition: The r th and s th product moment about the origin of the continuous random variables X and Y , denoted by $\mu'_{r,s}$, for $r = 1, 2, \dots$; $s = 1, 2, \dots$, is given by

$$\mu'_{r,s} = E[X^r Y^s] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^r y^s f(x, y) dx dy$$

Remarks:

- If $r = s = 1$, we have $\mu'_{1,1} = E[XY]$
- *Cauchy-Schwarz Inequality:* For any two random variables X and Y , we have $|E[XY]| \leq E[X^2]^{1/2} E[Y^2]^{1/2}$ provided that $E[|XY|]$ is finite.
- If X and Y are independent random variables, $E[h(X)v(Y)] = E(h(X))E(v(Y))$ for any two functions $h(X)$ and $v(Y)$. **[Warning:** The reverse is not true.]
- If X_1, X_2, \dots, X_n are independent random variables independent, $E[X_1 X_2 \dots X_n] = E(X_1) E(X_2) \dots E(X_n)$. **[Warning:** The reverse is not true.]

3.15 Product moments about the mean

Definition: The r th and s th product moment about the mean of the discrete random variables X and Y , denoted by $\mu_{r,s}$ is the expected value of $(X - \mu_X)^r (Y - \mu_Y)^s$, for $r = 1, 2, \dots$; $s = 1, 2, \dots$ which is given by

$$\mu_{r,s} = E[(X - \mu_X)^r (Y - \mu_Y)^s] = \sum_{(x,y) \in D_{(X,Y)}} (x - \mu_X)^r (y - \mu_Y)^s f_{X,Y}(x, y)$$

Remark: In the case $D_{(X,Y)}$ is countably infinite we have

$$\mu_{r,s} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (x_i - \mu_X)^r (y_j - \mu_Y)^s f_{X,Y}(x_i, y_j)$$

Definition: The r th and s th product moment about the mean of the continuous random variables X and Y , denoted by $\mu_{r,s}$, for $r = 1, 2, \dots$; $s = 1, 2, \dots$ is given by

$$\mu_{r,s} = E[(X - \mu_X)^r (Y - \mu_Y)^s] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)^r (y - \mu_Y)^s f(x, y) dx dy$$

3.16 The covariance

The covariance is a measure of the joint variability of two random variables. Formally it is defined as

$$\text{Cov}(X, Y) = \sigma_{XY} = \mu_{1,1} = E[(X - \mu_X)(Y - \mu_Y)]$$

If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the lesser values, i.e., the variables tend to show similar behavior, the covariance is positive. In the opposite case, when the greater values of one variable mainly correspond to the lesser values of the other, i.e., the variables tend to show opposite behavior, the covariance is negative.

Properties:

- $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.
- If X and Y are independent $\text{Cov}(X, Y) = 0$.
- If $Y = bZ$, where b is constant,

$$\text{Cov}(X, Y) = b\text{Cov}(X, Z).$$

- If $Y = V + W$,

$$\text{Cov}(X, Y) = \text{Cov}(X, V) + \text{Cov}(X, W).$$

- If $Y = b$, where b is constant,

$$\text{Cov}(X, Y) = 0.$$

- It follows from the Cauchy-Schwarz Inequality that $|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$.

3.17 The correlation coefficient

The covariance has the inconvenient of depending on the scale of both random variables. In order to obtain a measure of the joint variability of two random variables that do not depend on the scale we need to divide the covariance by the standard deviation of each random variable. This leads to the correlation coefficient [or Pearson correlation coefficient (Karl Pearson, 1857-1936)]

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Properties:

- It follows from the Cauchy-Schwarz Inequality that $-1 \leq \rho_{X,Y} \leq 1$.

If $Y = bX + a$, where b and a are constants

- $\rho_{X,Y} = 1$ if $b > 0$.
- $\rho_{X,Y} = -1$ if $b < 0$.
- $\rho_{X,Y} = 0$ if $b = 0$.

Exercise: Let X and Y be the random variables representing the population of monthly wages of husbands and wives in a particular community. Say, there are only three possible monthly wages in euros: 0, 1000, 2000. The joint probability distribution is

X	0	1000	2000
0	0.05	0.15	0.10
1000	0.10	0.10	0.30
2000	0.05	0.05	0.10

Compute $\rho_{X,Y}$.

Exercise: Let (X, Y) be a continuous two-dimensional random variables. with joint probability density function :

$$f_{X,Y}(x, y) = 4xy, \text{ for } 0 < x < 1 \text{ and } 0 < y < 1.$$

Compute $\rho_{X,Y}$.

3.18 Moments of linear functions of random variables

Summary of important results:

- If $Y = V + W$,

$$Var(Y) = Var(V) + Var(W) + 2Cov(V, W).$$

- If $Y = V - W$,

$$Var(Y) = Var(V) + Var(W) - 2Cov(V, W).$$

- If X_1, \dots, X_n are random variables and a_1, \dots, a_n are constants and $Y = \sum_{i=1}^n a_i X_i$, then

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

- If X_1, \dots, X_n are independent random variables and a_1, \dots, a_n are constants

$$Var(Y) = \sum_{i=1}^n a_i^2 Var(X_i).$$

- If X_1, \dots, X_n are random variables, a_1, \dots, a_n are constants and b_1, \dots, b_n are constants, $Y_1 = \sum_{i=1}^n a_i X_i$, and $Y_2 = \sum_{i=1}^n b_i X_i$ then

$$Cov(Y_1, Y_2) = \sum_{i=1}^n a_i b_i Var(X_i) + \sum_{i=1}^n \sum_{j=1, j < i}^n (a_i b_j + a_j b_i) Cov(X_i, X_j)$$

- If X_1, \dots, X_n are independent random variables, a_1, \dots, a_n are constants and b_1, \dots, b_n are constants, $Y_1 = \sum_{i=1}^n a_i X_i$, and $Y_2 = \sum_{i=1}^n b_i X_i$ then

$$Cov(Y_1, Y_2) = \sum_{i=1}^n a_i b_i Var(X_i).$$

3.19 Conditional expectations

Definition: If (X, Y) are discrete random variables, D_Y is the set of discontinuity points of $F_Y(y)$, $f_{Y|X=x}(y)$ is the value of the conditional probability function of Y given $X = x$ at y , and $u(Y, X)$ is a function of Y and X , the conditional expectation of $u(Y, X)$ given $X = x$, is given by

$$E[u(Y, X)|X = x] = \sum_{y \in D_Y} u(y, x) f_{Y|X=x}(y)$$

provided that it is finite.

Remark: If D_Y is countably infinite we have $D_Y = \{y_1, y_2, y_3, \dots\}$ and consequently

$$E[u(Y, X)|X = x] = \sum_{i=1}^{\infty} u(y_i, x) f_{Y|X=x}(y_i)$$

Definition: If (X, Y) are continuous random variables, $f_{Y|X=x}(y)$ is the value of the conditional probability density function of Y given $X = x$ at y , and $u(Y, X)$ is a function of Y and X , the conditional expectation of $u(Y, X)$ given $X = x$, is given by

$$E[u(Y, X)|X = x] = \int_{-\infty}^{+\infty} u(y, x) f_{Y|X=x}(y) dy$$

provided that it is finite.

Remarks:

1. The calculation of $E[u(Y, X)|X = x]$ requires that we fix the value of X at x . If x varies, the $E[u(Y, X)|X = x]$ represents a function of x , that is $g(x) = E[u(Y, X)|X = x]$. In this case we can consider $g(X) = E[u(Y, X)|X]$ a random variable.
2. If $u(Y, X) = Y$ we have the *conditional mean* of Y , $E[u(Y, X)|X = x] = E[Y|X = x] = \mu_{Y|x}$ (notice that this is a function of x).
3. Let $u(Y, X) = (Y - \mu_{Y|x})^2$ we have the *conditional variance* of Y

$$\begin{aligned} E[u(Y, X)|X = x] &= E\left[(Y - \mu_{Y|x})^2 | X = x\right] \\ &= E\left[(Y - E[u(Y)|X = x])^2 | X = x\right] \\ &= \text{Var}[Y|X = x] \end{aligned}$$

4. $\text{Var}[Y|X = x] = E[Y^2|X = x] - E[Y|X = x]^2$.
5. If Y and X are independent $E(Y|X = x) = E(Y)$.
6. Of course we can reverse the roles of Y and X , that is we can compute $E(u(X, Y)|Y = y)$, using definitions similar to those above. For instance for the continuous case we have $E(u(X, Y)|Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y=y}(x) dx$.

Exercise: Suppose the joint probability function of (X, Y) is given in the following table:

	X	0	1000	2000	$\mathcal{P}(Y = y)$
Y					
0		0.05	0.15	0.10	0.30
1000		0.10	0.10	0.30	0.50
2000		0.05	0.05	0.10	0.20
$\mathcal{P}(X = x)$		0.20	0.30	0.50	1

Compute $E(Y|X = 0)$.

Solution:

$$\begin{aligned}\mathcal{P}(Y = 0|X = 0) &= \frac{\mathcal{P}(Y = 0, X = 0)}{\mathcal{P}(X = 0)} = \frac{0.05}{0.2} = 0.25 \\ \mathcal{P}(Y = 1000|X = 0) &= \frac{\mathcal{P}(Y = 1000, X = 0)}{\mathcal{P}(X = 0)} = \frac{0.1}{0.2} = 0.5. \\ \mathcal{P}(Y = 2000|X = 0) &= \frac{\mathcal{P}(Y = 2000, X = 0)}{\mathcal{P}(X = 0)} = \frac{0.05}{0.2} = 0.25.\end{aligned}$$

and hence

$$\begin{aligned}E[Y|X = 0] &= 0 \times \underbrace{\mathcal{P}(Y = 0|X = 0)}_{0.25} \\ &+ 1000 \times \underbrace{\mathcal{P}(Y = 1000|X = 0)}_{0.5} + 2000 \times \underbrace{\mathcal{P}(Y = 2000|X = 0)}_{0.25} = 1000.\end{aligned}$$

Example: Let (X, Y) be a two-dimensional random variable with joint probability density function: $f_{X,Y}(x, y) = 15xy^2$ for $0 < y < x$, $0 < x < 1$ and $f_{X,Y}(x, y) = 0$ otherwise. Compute $E[Y|X = 0.75]$.

Solution: In this case $f_X(x) = 5x^4$, for $0 < x < 1$ and hence $f_{Y|X=x}(y) = \frac{3y^2}{x^3}$, $0 < y < x$. Therefore $E[Y|X = x] = \frac{3x}{4}$, for $0 < x < 1$ and consequently $E[Y|X = 0.75] = \frac{3(0.75)}{4} = 0.5625$.

3.20 The law of iterated expectations

Theorem (*Law of iterated Expectations*) $E(u(Y, X)) = E(E[u(Y, X)|X])$ provided that $E(|u(Y, X)|)$ is finite.

Remark: This theorem shows that there are two ways to compute $E(u(Y, X))$. The first is the direct way. The second way is to consider the following steps:

1. compute $E[u(Y, X)|X = x]$ and notice that this is a function solely of x that is we can write $g(x) = E[u(Y, X)|X = x]$,
2. according to the theorem replacing $g(x)$ by $g(X)$ and taking the mean we obtain $E[g(X)] = E[u(Y, X)]$ for this specific form of $g(X)$.

Example (cont): Let (X, Y) be a two-dimensional random variable with joint probability density function: $f_{X,Y}(x, y) = 15xy^2$ for $0 < y < x$, $0 < x < 1$ and $f_{X,Y}(x, y) = 0$ otherwise. Compute $E[Y]$ using the direct way and law of iterated expectations.

Solution: Notice that we have $f_Y(y) = \frac{15}{2}(y^2 - y^4)$, for $0 < y < 1$ hence

$$E(Y) = \int_0^1 y f_Y(y) dy = \frac{5}{8}$$

Now we use the law of the iterated expectations. We showed before that $E[Y|X = x] = \frac{3x}{4} = g(x)$. Hence according to the law of iterated expectation

$$\begin{aligned} E(Y) &= E(g(X)) \\ &= E\left(\frac{3X}{4}\right) = \frac{3}{4}E(X) \end{aligned}$$

But

$$E(X) = \int_0^1 x f_X(x) dx = \frac{5}{6}$$

Therefore $E(Y) = \frac{3}{4}E(X) = \frac{3}{4} \cdot \frac{5}{6} = \frac{5}{8}$.

Remarks: This theorem is useful in practice in the calculation of $E(u(Y, X))$ if we know $f_{Y|X=x}(y)$ or $E[u(Y, X)|X = x]$ and $f_X(x)$ (or some moments of X), but not $f_{X,Y}(x, y)$.

Exercise: Suppose we have two random variables Y and X such that

$$E(Y|X = x) = 5x$$

for all x and $E(X) = 0$, $E(X^2) = 0.2$. Use the law of iterated expectations to find the value of $cov(X, Y)$.

Solution: Notice that $cov(X, Y) = E(XY) - E(X)E(Y) = E(XY)$ as $E(X) = 0$. Let $u(Y, X) = XY$. By the law of iterated expectations $E(u(Y, X)) = E(E(u(Y, X)|X))$

Now notice that $E(u(Y, X)|X = x) = E(XY|X = x) = E(xY|X = x) = xE(Y|X = x) = 5x^2 = g(x)$. According to the law of iterated expectations $E(XY) = E(g(X)) = E(5X^2) = 5E(X^2) = 1$.