

Econ 205

Multivariate (Joint) Probability Distributions – 3

Course Outline

- **Probability**

- Probability and Inference
- The Discrete Case
- Calculating the Probability of an Event
- Conditional Probability and the Independence of Events
- Laws of Probability
- The Law of Total Probability and Bayes' Rule
- Introduction to Random Variables and their distributions

- **Discrete Random Variables and Their Probability Distributions**

- Discrete Random Variable
- The Probability Distribution for Discrete Random Variable
- The Expected Value of Random Variable
- Moments and Moment-Generating Functions
- The Bernoulli Probability Distribution
- The Binomial Probability Distribution
- The Geometric Probability Distribution
- The Negative Binomial Probability Distribution
- The Hypergeometric Probability Distribution
- The Poisson Probability Distribution
- Probability-Generating Functions

- **Continuous Random Variables and Their Probability Densities**

- Continuous Random Variable
- The Probability Distribution for Continuous Random Variable
- The Expected Value for Continuous Random Variable
- Moments and Moment-Generating Functions
- The Uniform Probability Distribution
- The Normal Probability Distribution
- Exponential Probability Distribution
- Chi-Square Probability Distribution
- Tchebysheff's Theorem

Course Outline (cont'd)

- **Multivariate Probability Distributions**

- Bivariate and Multivariate Probability Distributions
- Independent Random Variables
- Marginal and Conditional Probability Distributions
- The Expected Value of a Function of Random Variables
- Special Theorems
- The Covariance of Two Random Variables
- The Expected Value and Variance of Linear Functions of Random Variables
- The Multinomial Probability Distribution
- The Bivariate Normal Distribution
- Conditional Expectations

- **Functions of Random Variables**

- Finding the Probability Distribution of a Function of Random Variables
- The Method of Distribution Functions
- The Method of Transformations
- The Method of Moment-Generating Functions
- Multivariable Transformations Using Jacobians

Table of Contents of Devore & Berk

2	Probability	50
	Introduction	50
2.1	Sample Spaces and Events	51
2.2	Axioms, Interpretations, and Properties of Probability	56
2.3	Counting Techniques	66
2.4	Conditional Probability	74
2.5	Independence	84
3	Discrete Random Variables and Probability Distributions	96
	Introduction	96
3.1	Random Variables	97
3.2	Probability Distributions for Discrete Random Variables	101
3.3	Expected Values of Discrete Random Variables	112
3.4	Moments and Moment Generating Functions	121
3.5	The Binomial Probability Distribution	128
3.6	Hypergeometric and Negative Binomial Distributions	138
3.7	The Poisson Probability Distribution	146
4	Continuous Random Variables and Probability Distributions	158
	Introduction	158
4.1	Probability Density Functions and Cumulative Distribution Functions	159
4.2	Expected Values and Moment Generating Functions	171
4.3	The Normal Distribution	179
4.4	The Gamma Distribution and Its Relatives	194
4.5	Other Continuous Distributions	202
4.6	Probability Plots	210
4.7	Transformations of a Random Variable	220
5	Joint Probability Distributions	232
	Introduction	232
5.1	Jointly Distributed Random Variables	233
5.2	Expected Values, Covariance, and Correlation	245
5.3	Conditional Distributions	253
5.4	Transformations of Random Variables	265
5.5	Order Statistics	271

Content of this presentation

- Conditional Distributions
 - **Conditional Pmf** of Y given $X=x$ and **Conditional Pdf** of Y given $X=x$
 - **Conditional Mean** of Y given $X=x$ and **Conditional Variance** of Y given $X=x$
 - The Mean and Variance via the Conditional Mean and Conditional Variance: Two Theorems
 - Independence
- Bivariate Normal Distribution
- Examples

Conditional pmf of Y given X=x

Conditional pdf of Y given X=x

Definition:

Let X and Y be two discrete random variables with joint pmf $p(x,y)$ and marginal X pmf $p_X(x)$. Then for any x value such that $p_X(x) > 0$, the **conditional probability mass function of Y given $X = x$** is

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$$

An analogous formula holds in the continuous case. Let X and Y be two continuous random variables with joint pdf $f(x,y)$ and marginal X pdf $f_X(x)$. Then for any x value such that $f_X(x) > 0$, the **conditional probability density function of Y given $X = x$** is

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

Recall Marginal pmf s and Marginal pdf s:

The **marginal probability mass functions** of X and of Y , denoted by $p_X(x)$ and $p_Y(y)$, respectively, are given by

$$p_X(x) = \sum_y p(x,y) \quad p_Y(y) = \sum_x p(x,y)$$

The **marginal probability density functions** of X and Y , denoted by $f_X(x)$ and $f_Y(y)$, respectively, are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy \quad \text{for } -\infty < x < \infty$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y) dx \quad \text{for } -\infty < y < \infty$$

Remark: $f_i(y_i)$ is an alternative notation for $f_{y_i}(y_i)$.

Conditional Mean and Conditional Variance

Definition:

Let X and Y be two discrete random variables with conditional probability mass function $p_{Y|X}(y|x)$. Then the **conditional mean** or **expected value of Y given that $X = x$** is

$$\mu_{Y|X=x} = E(Y|X=x) = \sum_{y \in D_Y} y p_{Y|X}(y|x)$$

An analogous formula holds in the continuous case. Let X and Y be two continuous random variables with conditional probability density function $f_{Y|X}(y|x)$. Then

$$\mu_{Y|X=x} = E(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy$$

The conditional mean of any function $g(Y)$ can be obtained similarly. In the discrete case,

$$E(g(Y)|X=x) = \sum_{y \in D_Y} g(y) p_{Y|X}(y|x)$$

In the continuous case

$$E(g(Y)|X=x) = \int_{-\infty}^{\infty} g(y) f_{Y|X}(y|x) dy$$

The **conditional variance of Y given $X = x$** is

$$\sigma_{Y|X=x}^2 = V(Y|X=x) = E\{[Y - E(Y|X=x)]^2 | X=x\}$$

$$V(Y_1 | Y_2 = y_2) = E(Y_1^2 | Y_2 = y_2) - [E(Y_1 | Y_2 = y_2)]^2.$$

Compare with the Mean:

For a discrete r.v. Y :

$$E(Y) = \sum_y y p(y)$$

For a continuous r.v. Y :

$$E(Y) = \int_{-\infty}^{\infty} y f(y) dy,$$

provided that the integral exists.

Let $g(Y)$ be a function of Y .

For a discrete r.v. Y :

$$E[g(Y)] = \sum_{\text{all } y} g(y) p(y)$$

For a continuous r.v. Y :

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y) f(y) dy,$$

provided that the integral exists.

Compare with the Variance:

Let Y be a discrete random variable with probability function $p(y)$ and mean $E(Y) = \mu$; then

$$V(Y) = \sigma^2 = E[(Y - \mu)^2] = E(Y^2) - \mu^2.$$

The Mean and Variance via the Conditional Mean and Conditional Variance : Two Theorems

THEOREM 10.1

Let Y_1 and Y_2 denote random variables. Then

$$E(Y_1) = E[E(Y_1 | Y_2)],$$

where on the right-hand side the inside expectation is with respect to the conditional distribution of Y_1 given Y_2 and the outside expectation is with respect to the distribution of Y_2 .

Proof

Suppose that Y_1 and Y_2 are jointly continuous with joint density function $f(y_1, y_2)$ and marginal densities $f_1(y_1)$ and $f_2(y_2)$, respectively. Then

$$\begin{aligned} E(Y_1) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 f(y_1, y_2) dy_1 dy_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 f(y_1 | y_2) f_2(y_2) dy_1 dy_2 \\ &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} y_1 f(y_1 | y_2) dy_1 \right\} f_2(y_2) dy_2 \\ &= \int_{-\infty}^{\infty} E(Y_1 | Y_2 = y_2) f_2(y_2) dy_2 = E[E(Y_1 | Y_2)]. \end{aligned}$$

The proof is similar for the discrete case.

THEOREM 10.2

Let Y_1 and Y_2 denote random variables. Then

$$V(Y_1) = E[V(Y_1 | Y_2)] + V[E(Y_1 | Y_2)].$$

Proof

As previously indicated, $V(Y_1 | Y_2)$ is given by

$$V(Y_1 | Y_2) = E(Y_1^2 | Y_2) - [E(Y_1 | Y_2)]^2$$

and

$$E[V(Y_1 | Y_2)] = E[E(Y_1^2 | Y_2)] - E\{[E(Y_1 | Y_2)]^2\}.$$

By definition,

$$V[E(Y_1 | Y_2)] = E\{[E(Y_1 | Y_2)]^2\} - \{E[E(Y_1 | Y_2)]\}^2.$$

The variance of Y_1 is

$$\begin{aligned} V(Y_1) &= E[Y_1^2] - [E(Y_1)]^2 \\ &= E\{E[Y_1^2 | Y_2]\} - \{E[E(Y_1 | Y_2)]\}^2 \\ &= E\{E[Y_1^2 | Y_2]\} - E\{[E(Y_1 | Y_2)]^2\} + E\{[E(Y_1 | Y_2)]^2\} \\ &\quad - \{E[E(Y_1 | Y_2)]\}^2 \\ &= E[V(Y_1 | Y_2)] + V[E(Y_1 | Y_2)]. \end{aligned}$$

Independence

suppose there is independence in the discrete case. Then

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)} = \frac{p_X(x)p_Y(y)}{p_X(x)} = p_Y(y)$$

That is, independence implies that the conditional distribution of Y is the same as the unconditional distribution. The implication works in the other direction, too. If

$$p_{Y|X}(y|x) = p_Y(y)$$

then

$$\frac{p(x,y)}{p_X(x)} = p_Y(y)$$

so

$$p(x,y) = p_X(x)p_Y(y)$$

and therefore X and Y are independent.

The continuous case is similar.

Recall: Independence of random variables

Two random variables X and Y are said to be **independent** if for every pair of x and y values,

$$p(x,y) = p_X(x) \cdot p_Y(y) \quad \text{when } X \text{ and } Y \text{ are discrete}$$

or

(5.1)

$$f(x,y) = f_X(x) \cdot f_Y(y) \quad \text{when } X \text{ and } Y \text{ are continuous}$$

If (5.1) is not satisfied for all (x, y) , then X and Y are said to be **dependent**.

The Bivariate Normal Distribution-1

The bivariate normal density is

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\{[(x-\mu_1)/\sigma_1]^2 - 2\rho(x-\mu_1)(y-\mu_2)/\sigma_1\sigma_2 + [(y-\mu_2)/\sigma_2]^2\}/[2(1-\rho^2)]}$$

There are five parameters, including the mean μ_1 and the standard deviation σ_1 of X and the mean μ_2 and the standard deviation σ_2 of Y . The fifth parameter ρ is the correlation between X and Y .

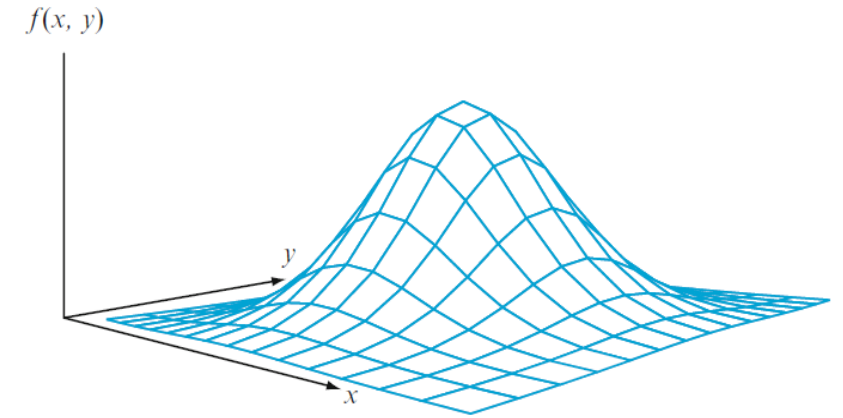
Note that for a bivariate normal distribution:

$\rho = 0$ if and only if the two rv's are independent.

(Note that this statement is not true in general.)

The marginal pdf of X is:

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\{[(x-\mu_1)/\sigma_1]^2\}/2}$$



A graph of the bivariate normal pdf

The Bivariate Normal Distribution-2

The conditional density $f_{Y|X}(y|x)$ is a normal density with mean and variance given by

$$\mu_{Y|X=x} = E(Y|X = x) = \mu_2 + \rho\sigma_2 \frac{x - \mu_1}{\sigma_1}$$

$$\sigma_{Y|X=x}^2 = V(Y|X = x) = \sigma_2^2(1 - \rho^2)$$

Observe the following:

When $\rho = 0$, the conditional mean is the mean of Y and the conditional variance is just the variance of Y . In other words, if $\rho = 0$, then the conditional distribution of Y is the same as the unconditional distribution of Y .

EXAMPLES

Example 1: Conditional pmf of Y given X=x

Assume that the joint distribution of X and Y is given in the following table:

$p(x, y)$		y		
		0	100	200
x	100	.20	.10	.20
	250	.05	.15	.30

The distribution of Y depends on X . In particular, let's find the conditional probability that Y is 200, given that X is 250, using the definition of conditional probability

$$P(Y = 200|X = 250) = \frac{P(Y = 200 \text{ and } X = 250)}{P(X = 250)} = \frac{.3}{.05 + .15 + .3} = .6$$

With our new definition we obtain the same result

$$p_{Y|X}(200|250) = \frac{p(250, 200)}{p_X(250)} = \frac{.3}{.05 + .15 + .3} = .6$$

Recall the definition of conditional pmf of Y given X=x:

Let X and Y be two discrete random variables with joint pmf $p(x,y)$ and marginal X pmf $p_X(x)$. Then for any x value such that $p_X(x) > 0$, the **conditional probability mass function of Y given X = x** is

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(x)}$$

Example 1 (cont'd):

The conditional probabilities for the two other possible values of Y are

$$p_{Y|X}(0|250) = \frac{p(250,0)}{p_X(250)} = \frac{.05}{.05 + .15 + .3} = .1$$
$$p_{Y|X}(100|250) = \frac{p(250,100)}{p_X(250)} = \frac{.15}{.05 + .15 + .3} = .3$$

Thus, $p_{Y|X}(0|250) + p_{Y|X}(100|250) + p_{Y|X}(200|250) = .1 + .3 + .6 = 1$. This is no coincidence; conditional probabilities satisfy the properties of ordinary probabilities. They are nonnegative and they sum to 1. Essentially, the denominator in the definition of conditional probability is designed to make the total be 1.

Reversing the roles of X and Y , we find the conditional probabilities for X , given that $Y = 0$:

$$p_{X|Y}(100|0) = \frac{p(100,0)}{p_Y(0)} = \frac{.20}{.20 + .05} = .8$$
$$p_{X|Y}(250|0) = \frac{p(250,0)}{p_Y(0)} = \frac{.05}{.20 + .05} = .2$$

Again, the conditional probabilities add to 1.

Recall the definition of conditional pmf of Y given $X=x$:

Let X and Y be two discrete random variables with joint pmf $p(x,y)$ and marginal X pmf $p_X(x)$. Then for any x value such that $p_X(x) > 0$, the **conditional probability mass function of Y given $X = x$** is

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}$$



Example 2: Conditional pdf of Y given X=x

Assume that the joint pdf of X and Y is

$$f(x,y) = \begin{cases} 24xy & 0 \leq x \leq 1, 0 \leq y \leq 1, x+y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

It can be shown that $f_X(x) = \begin{cases} 12x(1-x)^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$

The conditional pdf of Y given X=x is defined for $0 < x < 1$ and undefined otherwise. **For $0 < x < 1$:**

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{24xy}{12x(1-x)^2} = \frac{2y}{(1-x)^2} \quad 0 \leq y \leq 1-x$$

This can be used to get conditional probabilities for Y. For example,

$$P(Y \leq .25 | X = .5) = \int_{-\infty}^{.25} f_{Y|X}(y|.5) dy = \int_0^{.25} \frac{2y}{(1-.5)^2} dy = [4y^2]_0^{.25} = .25$$

Just as in the discrete case, the conditional distribution assigns a total probability of 1 to the set of all possible Y values. That is, integrating the conditional density over its set of possible values should yield 1:

$$\int_{-\infty}^{\infty} f_{Y|X}(y|x) dy = \int_0^{1-x} \frac{2y}{(1-x)^2} dy = \left[\frac{y^2}{(1-x)^2} \right]_0^{1-x} = 1$$

Whenever you calculate a conditional density, we recommend doing this integration as a validity check. ■

Recall the definition of conditional pdf of Y given X=x:

Let X and Y be two continuous random variables with joint pdf $f(x,y)$ and marginal X pdf $f_X(x)$. Then for any x value such that $f_X(x) > 0$, the **conditional probability density function of Y given $X = x$** is

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

Example 3: Conditional pdf of Y given X=x

EXAMPLE

A soft-drink machine has a random amount Y_2 in supply at the beginning of a given day and dispenses a random amount Y_1 during the day (with measurements in gallons). It is not resupplied during the day, and hence $Y_1 \leq Y_2$. It has been observed that Y_1 and Y_2 have a joint density given by

$$f(y_1, y_2) = \begin{cases} 1/2, & 0 \leq y_1 \leq y_2 \leq 2, \\ 0 & \text{elsewhere.} \end{cases}$$

That is, the points (y_1, y_2) are uniformly distributed over the triangle with the given boundaries. Find the conditional density of Y_1 given $Y_2 = y_2$. Evaluate the probability that less than 1/2 gallon will be sold, given that the machine contains 1.5 gallons at the start of the day.

Solution The marginal density of Y_2 is given by

$$f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) dy_1.$$

Thus,

$$f_2(y_2) = \begin{cases} \int_0^{y_2} (1/2) dy_1 = (1/2)y_2, & 0 \leq y_2 \leq 2, \\ \int_{-\infty}^{\infty} 0 dy_1 = 0, & \text{elsewhere.} \end{cases}$$

Note that $f_2(y_2) > 0$ if and only if $0 < y_2 \leq 2$. Thus, for any $0 < y_2 \leq 2$, using Definition

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} = \frac{1/2}{(1/2)(y_2)} = \frac{1}{y_2}, \quad 0 \leq y_1 \leq y_2.$$

Also, $f(y_1|y_2)$ is undefined if $y_2 \leq 0$ or $y_2 > 2$. The probability of interest is

$$P(Y_1 \leq 1/2 | Y_2 = 1.5) = \int_{-\infty}^{1/2} f(y_1|y_2 = 1.5) dy_1 = \int_0^{1/2} \frac{1}{1.5} dy_1 = \frac{1/2}{1.5} = \frac{1}{3}.$$

If the machine contains 2 gallons at the start of the day, then

$$P(Y_1 \leq 1/2 | Y_2 = 2) = \int_0^{1/2} \frac{1}{2} dy_1 = \frac{1}{4}.$$

Thus, the conditional probability that $Y_1 \leq 1/2$ given $Y_2 = y_2$ changes appreciably depending on the particular choice of y_2 . ■

Recall the definition of conditional pdf of Y given X=x:

Let X and Y be two continuous random variables with joint pdf $f(x,y)$ and marginal X pdf $f_X(x)$. Then for any x value such that $f_X(x) > 0$, the **conditional probability density function of Y given X = x** is

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

Example 4: Conditional Mean and Conditional Variance

Consider the joint pdf of Example 2 again. Let's find the conditional mean and the conditional variance. The conditional pdf of Y given $X=x$ for $0 < x < 1$ was computed as:

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{24xy}{12x(1-x)^2} = \frac{2y}{(1-x)^2} \quad 0 \leq y \leq 1-x$$

Using the definitions, for $0 < x < 1$

$$\begin{aligned} \mu_{Y|X=x} &= E(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_0^{1-x} y \frac{2y}{(1-x)^2} dy \\ &= \frac{2}{3}(1-x) \end{aligned}$$

$$E(Y^2|X=x) = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x) dy = \int_0^{1-x} y^2 \frac{2y}{(1-x)^2} dy = \frac{(1-x)^2}{2} \quad \text{for } 0 < x < 1$$

Then the conditional variance is

$$\sigma_{Y|X=x}^2 = V(Y|X=x) = E(Y^2|X=x) - \mu_{Y|X=x}^2 = \frac{(1-x)^2}{2} - \frac{4(1-x)^2}{9} = \frac{(1-x)^2}{18} \quad \text{for } 0 < x < 1$$

Recall the definition of conditional pdf of Y given $X=x$:

Let X and Y be two continuous random variables with joint pdf $f(x,y)$ and marginal X pdf $f_X(x)$. Then for any x value such that $f_X(x) > 0$, the **conditional probability density function of Y given $X = x$** is

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

Example 5: The Mean and Variance via the Conditional Mean and Conditional Variance

EXAMPLE A quality control plan for an assembly line involves sampling $n = 10$ finished items per day and counting Y , the number of defectives. If p denotes the probability of observing a defective, then Y has a binomial distribution, assuming that a large number of items are produced by the line. But p varies from day to day and is assumed to have a uniform distribution on the interval from 0 to $1/4$. Find the expected value of Y . Find the variance of Y .

Solution From Theorem 4.1, we know that $E(Y) = E[E(Y|p)]$. For a given p , Y has a binomial distribution, and hence $E(Y|p) = np$. Thus,

$$E(Y) = E[E(Y|p)] = E(np) = nE(p) = n \left(\frac{1/4 + 0}{2} \right) = \frac{n}{8},$$

and for $n = 10$

$$E(Y) = 10/8 = 1.25.$$

In the long run, this inspection policy will average 1.25 defectives per day.

From Theorem 4.2, we know that

$$V(Y) = E[V(Y|p)] + V[E(Y|p)].$$

For a given p , Y has a binomial distribution, and hence $E(Y|p) = np$ and $V(Y|p) = npq$. Thus,

$$\begin{aligned} V(Y) &= E[V(Y|p)] + V[E(Y|p)] \\ &= E(npq) + V(np) = nE[p(1-p)] + n^2V(p). \end{aligned}$$

Because p is uniformly distributed on the interval $(0, 1/4)$ and $E(p^2) = V(p) + [E(p)]^2$, it follows that

$$E(p) = \frac{1}{8}, \quad V(p) = \frac{(1/4 - 0)^2}{12} = \frac{1}{192}, \quad E(p^2) = \frac{1}{192} + \frac{1}{64} = \frac{1}{48}.$$

Thus,

$$\begin{aligned} V(Y) &= nE[p(1-p)] + n^2V(p) = n[E(p) - E(p^2)] + n^2V(p) \\ &= n\left(\frac{1}{8} - \frac{1}{48}\right) + n^2\left(\frac{1}{192}\right) = \frac{5n}{48} + \frac{n^2}{192}, \end{aligned}$$

and for $n = 10$,

$$V(Y) = 50/48 + 100/192 = 1.5625.$$

Recall the following two theorems:

Let Y_1 and Y_2 denote random variables. Then

$$E(Y_1) = E[E(Y_1 | Y_2)],$$

Let Y_1 and Y_2 denote random variables. Then

$$V(Y_1) = E[V(Y_1 | Y_2)] + V[E(Y_1 | Y_2)].$$

Example 6: The Bivariate Normal Distribution

Let X be mother's height and Y be daughter's height. A similar situation was one of the first applications of the bivariate normal distribution, by Francis Galton in 1886, and the data was found to fit the distribution very well. Suppose a bivariate normal distribution with mean $\mu_1 = 64$ in. and standard deviation $\sigma_1 = 3$ in. for X and mean $\mu_2 = 65$ in. and standard deviation $\sigma_2 = 3$ in. for Y . Here $\mu_2 > \mu_1$, which is in accord with the increase in height from one generation to the next. Assume $\rho = .4$. Then

$$\mu_{Y|X=x} = \mu_2 + \rho\sigma_2 \frac{x - \mu_1}{\sigma_1} = 65 + .4(3) \frac{x - 64}{3} = 65 + .4(x - 64) = .4x + 39.4$$

$$\sigma_{Y|X=x}^2 = V(Y|X = x) = \sigma_2^2(1 - \rho^2) = 9(1 - .4^2) = 7.56 \text{ and } \sigma_{Y|X=x} = 2.75$$

Notice that the conditional variance is 16% less than the variance of Y . Squaring the correlation gives the percentage by which the conditional variance is reduced relative to the variance of Y . ■

Recall:

The conditional density $f_{Y|X}(y|x)$ is a normal density with mean and variance given by

$$\mu_{Y|X=x} = E(Y|X = x) = \mu_2 + \rho\sigma_2 \frac{x - \mu_1}{\sigma_1}$$

$$\sigma_{Y|X=x}^2 = V(Y|X = x) = \sigma_2^2(1 - \rho^2)$$

References

- Devore, J.L. and Berk, K.N. (2012) Modern Mathematical Statistics with Applications, 2nd Edition, New York: Springer.
- Wackerly, D.D., Mendenhall, W. and Scheaffer, R.L., (2008) Mathematical Statistics With Applications, 7th Edition, Pacific Grove: Duxbury.