

Econ 205

Multivariate (Joint) Probability Distributions – 4

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

Content of this presentation

- The Multinomial Probability Distribution
 - Definitions
 - Theorem
 - Example

The Multinomial Probability Distribution: Definitions

A multinomial experiment is a generalization of the binomial experiment.

DEFINITION

A *multinomial experiment* possesses the following properties:

1. The experiment consists of n identical trials.
2. The outcome of each trial falls into one of k classes or cells.
3. The probability that the outcome of a single trial falls into cell i , is p_i , $i = 1, 2, \dots, k$ and remains the same from trial to trial. Notice that $p_1 + p_2 + p_3 + \dots + p_k = 1$.
4. The trials are independent.
5. The random variables of interest are $Y_1, Y_2, \dots, Y_k$, where Y_i equals the number of trials for which the outcome falls into cell i . Notice that $Y_1 + Y_2 + Y_3 + \dots + Y_k = n$.

DEFINITION

Assume that $p_1, p_2, \dots, p_k$ are such that $\sum_{i=1}^k p_i = 1$, and $p_i > 0$ for $i = 1, 2, \dots, k$. The random variables $Y_1, Y_2, \dots, Y_k$, are said to have a *multinomial distribution* with parameters n and $p_1, p_2, \dots, p_k$ if the joint probability function of $Y_1, Y_2, \dots, Y_k$ is given by

$$p(y_1, y_2, \dots, y_k) = \frac{n!}{y_1! y_2! \dots y_k!} p_1^{y_1} p_2^{y_2} \dots p_k^{y_k},$$

where, for each i , $y_i = 0, 1, 2, \dots, n$ and $\sum_{i=1}^k y_i = n$.

The Multinomial Probability Distribution: Theorem

THEOREM

If $Y_1, Y_2, \dots, Y_k$ have a multinomial distribution with parameters n and $p_1, p_2, \dots, p_k$, then

1. $E(Y_i) = np_i, V(Y_i) = np_i q_i$.
2. $\text{Cov}(Y_s, Y_t) = -np_s p_t$, if $s \neq t$.

Proof

The marginal distribution of Y_i can be used to derive the mean and variance. Recall that Y_i may be interpreted as the number of trials falling into cell i . Imagine all of the cells, excluding cell i , combined into a single large cell. Then every trial will result in cell i or in a cell other than cell i , with probabilities p_i and $1 - p_i$, respectively. Thus, Y_i possesses a binomial marginal probability distribution. Consequently,

$$E(Y_i) = np_i \quad \text{and} \quad V(Y_i) = np_i q_i, \quad \text{where } q_i = 1 - p_i.$$

The same results can be obtained by setting up the expectations and evaluating. For example,

$$E(Y_1) = \sum_{y_1} \sum_{y_2} \cdots \sum_{y_k} y_1 \frac{n!}{y_1! y_2! \cdots y_k!} p_1^{y_1} p_2^{y_2} \cdots p_k^{y_k}.$$

Because we have already derived the expected value and variance of Y_i , we leave the summation of this expectation to the interested reader.

Think of the multinomial experiment as a sequence of n independent trials and define, for $s \neq t$,

$$U_i = \begin{cases} 1, & \text{if trial } i \text{ results in class } s, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$W_i = \begin{cases} 1, & \text{if trial } i \text{ results in class } t, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$Y_s = \sum_{i=1}^n U_i \quad \text{and} \quad Y_t = \sum_{j=1}^n W_j.$$

The Multinomial Probability Distribution: Theorem (cont'd)

(Because $U_i = 1$ or 0 depending upon whether the i th trial resulted in class s , Y_s is simply the sum of a series of 0s and 1s. A 1 occurs in the sum everytime we observe an item from class s , and a 0 occurs everytime we observe any other class. Thus, Y_s is simply the number of times class s is observed. A similar interpretation applies to Y_t .)

Notice that U_i and W_i cannot both equal 1 (the i th item cannot simultaneously be in classes s and t). Thus, the product $U_i W_i$ always equals zero, and $E(U_i W_i) = 0$. The following results allow us to evaluate $\text{Cov}(Y_s, Y_t)$:

$$E(U_i) = p_s$$

$$E(W_j) = p_t$$

$$\text{Cov}(U_i, W_j) = 0, \quad \text{if } i \neq j \text{ because the trials are independent}$$

$$\text{Cov}(U_i, W_i) = E(U_i W_i) - E(U_i)E(W_i) = 0 - p_s p_t$$

From Theorem 5.12, we then have

$$\begin{aligned} \text{Cov}(Y_s, Y_t) &= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(U_i, W_j) \\ &= \sum_{i=1}^n \text{Cov}(U_i, W_i) + \sum_{i \neq j} \text{Cov}(U_i, W_j) \\ &= \sum_{i=1}^n (-p_s p_t) + \sum_{i \neq j} 0 = -n p_s p_t. \end{aligned}$$

The covariance here is negative, which is to be expected because a large number of outcomes in cell s would force the number in cell t to be small.

Recall another theorem from earlier:

Let $Y_1, Y_2, \dots, Y_n$ and $X_1, X_2, \dots, X_m$ be random variables with $E(Y_i) = \mu_i$ and $E(X_j) = \xi_j$. Define

$$U_1 = \sum_{i=1}^n a_i Y_i \quad \text{and} \quad U_2 = \sum_{j=1}^m b_j X_j$$

for constants $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_m$. Then the following hold:

- a $E(U_1) = \sum_{i=1}^n a_i \mu_i$.
- b $V(U_1) = \sum_{i=1}^n a_i^2 V(Y_i) + 2 \sum \sum_{1 \leq i < j \leq n} a_i a_j \text{Cov}(Y_i, Y_j)$, where the double sum is over all pairs (i, j) with $i < j$.
- c $\text{Cov}(U_1, U_2) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(Y_i, X_j)$.

Example:

EXAMPLE

According to recent census figures, the proportions of adults (persons over 18 years of age) in the United States associated with five age categories are as given in the following table.

Age	Proportion
18–24	.18
25–34	.23
35–44	.16
45–64	.27
65+	.16

If these figures are accurate and five adults are randomly sampled, find the probability that the sample contains one person between the ages of 18 and 24, two between the ages of 25 and 34, and two between the ages of 45 and 64.

Solution We will number the five age classes 1, 2, 3, 4, and 5 from top to bottom and will assume that the proportions given are the probabilities associated with each of the classes. Then we wish to find

$$p(y_1, y_2, y_3, y_4, y_5) = \frac{n!}{y_1! y_2! y_3! y_4! y_5!} p_1^{y_1} p_2^{y_2} p_3^{y_3} p_4^{y_4} p_5^{y_5},$$

for $n = 5$ and $y_1 = 1, y_2 = 2, y_3 = 0, y_4 = 2$, and $y_5 = 0$. Substituting these values into the formula for the joint probability function, we obtain

$$\begin{aligned} p(1, 2, 0, 2, 0) &= \frac{5!}{1! 2! 0! 2! 0!} (.18)^1 (.23)^2 (.16)^0 (.27)^2 (.16)^0 \\ &= 30(.18)(.23)^2 (.27)^2 = .0208. \end{aligned}$$



Reference

- Wackerly, D.D., Mendenhall, W. and Scheaffer, R.L., (2008)
Mathematical Statistics With Applications, 7th Edition, Pacific Grove:
Duxbury.