

Random Variables and Applications

OPRE 6301

Random Variables...

As noted earlier, *variability* is omnipresent in the business world. To model variability probabilistically, we need the concept of a **random variable**.

A random variable is a *numerically* valued variable which takes on different values with given probabilities.

Examples:

The return on an investment in a one-year period

The price of an equity

The number of customers entering a store

The sales volume of a store on a particular day

The turnover rate at your organization next year

Types of Random Variables...

Discrete Random Variable:

— one that takes on a *countable* number of possible values, e.g.,

- total of roll of two dice: 2, 3, ..., 12
- number of desktops sold: 0, 1, ...
- customer count: 0, 1, ...

Continuous Random Variable:

— one that takes on an *uncountable* number of possible values, e.g.,

- interest rate: 3.25%, 6.125%, ...
- task completion time: a nonnegative value
- price of a stock: a nonnegative value

Basic Concept: Integer or rational numbers are discrete, while real numbers are continuous.

Probability Distributions. . .

“Randomness” of a random variable is described by a **probability distribution**. Informally, the probability distribution specifies the probability or likelihood for a random variable to assume a particular value.

Formally, let X be a random variable and let x be a possible value of X . Then, we have two cases.

Discrete: the probability mass function of X specifies $P(x) \equiv P(X = x)$ for all possible values of x .

Continuous: the probability density function of X is a function $f(x)$ that is such that $f(x) \cdot h \approx P(x < X \leq x + h)$ for small positive h .

Basic Concept: The probability mass function specifies the actual probability, while the probability density function specifies the probability rate; both can be viewed as a measure of “likelihood.”

The continuous case will be discussed in Chapter 8.

Discrete Distributions...

A probability mass function must satisfy the following two requirements:

1. $0 \leq P(x) \leq 1$ for all x
2. $\sum_{\text{all } x} P(x) = 1$

Empirical data can be used to estimate the probability mass function. Consider, for example, the number of TVs in a household...

No. of TVs	No. of Households	x	$P(x)$
0	1,218	0	0.012
1	32,379	1	0.319
2	37,961	2	0.374
3	19,387	3	0.191
4	7,714	4	0.076
5	2,842	5	0.028
	101,501		1.000

For $x = 0$, the probability 0.012 comes from $1,218/101,501$. Other probabilities are estimated similarly.

Properties of Discrete Distributions. . .

Realized values of a discrete random variable can be viewed as samples from a conceptual/theoretical **population**.

For example, suppose a household is randomly drawn, or sampled, from the population *governed* by the probability mass function specified in the previous table. What is the probability for us to observe the event $\{X = 3\}$? Answer: 0.191. That X turns out to be 3 in a random sample is called a realization. Similarly, the realization $X = 2$ has probability 0.374.

We can therefore compute the population mean, variance, and so on. Results of such calculations are examples of population **parameters**.

Details. . .

Population Mean — Expected Value...

The population mean is the **weighted average** of all of its values. The weights are specified by the probability mass function. This parameter is also called the **expected value** of X and is denoted by $E(X)$.

The formal definition is similar to computing sample mean for grouped data:

$$\mu = E(X) \equiv \sum_{\text{all } x} x P(x). \quad (1)$$

Example: Expected No. of TVs

Let X be the number of TVs in a household.

Then,

$$\begin{aligned} E(X) &= 0 \cdot 0.012 + 1 \cdot 0.319 + \cdots + 5 \cdot 0.028 \\ &= 2.084 \end{aligned}$$

The Excel function SUMPRODUCT() can be used for this computation.

Interpretation

What does it mean when we say $E(X) = 2.084$ in the previous example? Do we “expect” to see any household to have 2.084 TVs?

The correct answer is that the expected value should be interpreted as a **long-run average**. Formally, let x_1, x_2, \dots, x_n be n (independent) realizations of X ; then, we *expect*:

$$E(X) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i.$$

Such a statement is called a *law of large numbers*.

Thus, in the previous example, the average number of TVs in a large number of randomly-selected households will approach the expected value 2.084.

Population Variance...

The population variance is calculated similarly. It is the weighted average of the **squared deviations** from the mean. Formally,

$$\sigma^2 = V(X) \equiv \sum_{\text{all } x} (x - \mu)^2 P(x). \quad (2)$$

Since (2) is an expected value (of $(X - \mu)^2$), it should be interpreted as the *long-run* average of squared deviations from the mean. Thus, the parameter σ^2 is a measure of the extent of variability in successive realizations of X .

Similar to sample variance, there is a “short-cut” formula:

$$\sigma^2 = V(X) = \sum_{\text{all } x} x^2 P(x) - \mu^2. \quad (3)$$

The standard deviation is given by:

$$\sigma = \sqrt{\sigma^2}. \quad (4)$$

Example: Variance of No. of TVs

Let X be the number of TVs in a household.

Then,

$$\begin{aligned} V(X) &= (0 - 2.084)^2 \cdot 0.012 + \cdots + (5 - 2.084)^2 \cdot 0.028 \\ &= 1.107; \end{aligned}$$

or,

$$\begin{aligned} V(X) &= 0^2 \cdot 0.012 + \cdots + 5^2 \cdot 0.028 - 2.084^2 \\ &= 1.107. \end{aligned}$$

Thus, on average, we expect X to have a squared deviation of 1.107 from the mean 2.084.

The standard deviation is: $\sigma = \sqrt{1.107} = 1.052$.

General Laws...

Expected-Value Calculations...

Calculations involving the expected value obey the following important laws:

1. $E(c) = c$

— the expected value of a constant (c) is just the value of the constant

2. $E(X + c) = E(X) + c$

— “translating” X by c has the same effect on the expected value; in other words, we can distribute an expected-value calculation into a sum

3. $E(cX) = cE(X)$

— “scaling” X by c has the same effect on the expected value; in other words, we can pull the constant c out of an expected-value calculation

Variance Calculations...

Calculations involving the variance obey the following important laws:

1. $V(c) = 0$

— the variance of a constant (c) is zero

2. $V(X + c) = V(X)$

— “translating” X by c has *no* effect on the variance

3. $V(cX) = c^2 V(X)$

— “scaling” X by c boosts the variance by a factor of c^2 ; in other words, when we pull out a constant c in a variance calculation, the constant should be squared (note however that the standard deviation of cX equals c times the standard deviation of X)

Example: Sales versus Profit

The monthly sales, X , of a company have a mean of \$25,000 and a standard deviation of \$4,000. Profits, Y , are calculated by multiplying sales by 0.3 and subtracting fixed costs of \$6,000.

What are the mean profit and the standard deviation of profit?

We know that:

$$E(X) = 25000,$$

$$V(X) = 4000^2 = 16000000, \text{ and}$$

$$Y = 0.3X - 6000.$$

Therefore,

$$\begin{aligned} E(Y) &= 0.3E(X) - 6000 \\ &= 0.3 \cdot 25000 - 6000 \\ &= 1500 \end{aligned}$$

and

$$\begin{aligned} \sigma &= \sqrt{0.3^2 V(X)} \\ &= \sqrt{0.09 \cdot 16000000} \\ &= 1200. \end{aligned}$$

Bivariate Distributions...

Up to now, we have looked at **univariate** distributions, i.e., probability distributions in *one* variable.

Bivariate distributions, also called **joint distributions**, are probabilities of combinations of *two* variables.

For discrete variables X and Y , the joint probability distribution or joint probability mass function of X **and** Y is defined as:

$$P(x, y) \equiv P(X = x \text{ and } Y = y)$$

for all *pairs* of values x and y .

As in the univariate case, we require:

1. $0 \leq P(x, y) \leq 1$ for all x and y
2. $\sum_{\text{all } x} \sum_{\text{all } y} P(x, y) = 1$

Example: Houses Sold by Two Agents

Mark and Lisa are two real estate agents. Let X and Y be the respective numbers of houses sold by them in a month. Based on past sales, we estimated the following joint probabilities for X and Y .

		X		
		0	1	2
Y	0	0.12	0.42	0.06
	1	0.21	0.06	0.03
	2	0.07	0.02	0.01

Thus, for example $P(0, 1) = 0.21$, meaning that the *joint* probability for Mark and Lisa to sell 0 and 1 houses, respectively, is 0.21. Other entries in the table are interpreted similarly.

Note that the sum of all entries must equal to 1.

Marginal Probabilities

In the previous example, the *marginal probabilities* are calculated by summing across rows and down columns:

		X			
		0	1	2	
Y	0	0.12	0.42	0.06	0.6
	1	0.21	0.06	0.03	0.3
	2	0.07	0.02	0.01	0.1
		0.4	0.5	0.1	1.0

This gives us the probability mass functions for X and Y *individually*:

X		Y	
x	$P(x)$	y	$P(y)$
0	0.4	0	0.6
1	0.5	1	0.3
2	0.1	2	0.1

Thus, for example, the *marginal* probability for Mark to sell 1 house is 0.5.

Independence

Two variables X and Y are said to be *independent* if

$$P(X = x \text{ and } Y = y) = P(X = x)P(Y = y)$$

for all x and y . That is, the joint probabilities equal the product of marginal probabilities. This is similar to the definition of independent *events*.

In the houses-sold example, we have

$$P(X = 0 \text{ and } Y = 2) = 0.07,$$

$$P(X = 0) = 0.4, \text{ and } P(Y = 2) = 0.1.$$

Hence, X and Y are *not* independent.

Properties of Bivariate Distributions...

Expected values, Variances, and Standard Deviations...

These *marginal* parameters are computed via earlier formulas.

Consider the previous example again. Then, for Mark, we have

$$E(X) = 0.7,$$

$$V(X) = 0.41, \text{ and}$$

$$\sigma_X = 0.64;$$

and for Lisa, we have

$$E(Y) = 0.5,$$

$$V(Y) = 0.45, \text{ and}$$

$$\sigma_Y = 0.67.$$

Covariance...

The **covariance** between two discrete variables is defined as:

$$COV(X, Y) \equiv \sum_{\text{all } x} \sum_{\text{all } y} (x - \mu_X)(y - \mu_Y)P(x, y). \quad (5)$$

This is equivalent to:

$$COV(X, Y) = \sum_{\text{all } x} \sum_{\text{all } y} xy P(x, y) - \mu_X \mu_Y. \quad (6)$$

Example: Houses Sold

$$\begin{aligned} COV(X, Y) &= (0 - 0.7)(0 - 0.5) \cdot 0.12 + \dots \\ &\quad + (2 - 0.7)(2 - 0.5) \cdot 0.01 \\ &= -0.15; \end{aligned}$$

or,

$$\begin{aligned} COV(X, Y) &= 0 \cdot 0 \cdot 0.12 + \dots + 2 \cdot 2 \cdot 0.01 \\ &\quad - 0.7 \cdot 0.5 \\ &= -0.15. \end{aligned}$$

Coefficient of Correlation...

As usual, the coefficient of correlation is given by:

$$\rho_{X,Y} \equiv \frac{COV(X, Y)}{\sigma_X \sigma_Y}. \quad (7)$$

Example: Houses Sold

$$\rho_{X,Y} = \frac{-0.15}{0.64 \cdot 0.67} = -0.35.$$

This indicates that there is a bit of *negative* relationship between the numbers of houses sold by Mark and Lisa.

Is this surprising?

Sum of Two Variables...

The bivariate distribution allows us to develop the probability distribution of the sum of two variables, which is of interest in many applications.

In the houses-sold example, we could be interested in the probability for having two houses sold (by either Mark or Lisa) in a month. This can be computed by adding the probabilities for all combinations of (x, y) pairs that result in a sum of 2:

$$P(X + Y = 2) = P(0, 2) + P(1, 1) + P(2, 0) = 0.19 .$$

Using this method, we can derive the probability mass function for the variable $X + Y$:

$x + y$	$P(x + y)$
0	0.12
1	0.63
2	0.19
3	0.05
4	0.01

The expected value and variance of $X + Y$ obey the following basic laws...

1. $E(X + Y) = E(X) + E(Y)$

2. $V(X + Y) = V(X) + V(Y) + 2COV(X, Y)$

If X and Y happens to be *independent*, then $COV(X, Y) = 0$ and thus

$$V(X + Y) = V(X) + V(Y).$$

Example: Houses Sold

$$E(X + Y) = 0.7 + 0.5 = 1.2,$$

$$V(X + Y) = 0.41 + 0.45 + 2(-0.15) = 0.56, \text{ and}$$

$$\sigma_{X+Y} = \sqrt{0.56} = 0.75.$$

Note that the negative correlation between X and Y had a *variance-reduction* effect on $X + Y$. This is an important concept. One application is that investing in both stocks and bonds could result in reduced variability or *risk*.

Business Applications...

Although our discussion may have seemed somewhat on the theoretical side, it turns out that some of the most applicable aspects of probability theory to business problems are through the concept of random variables.

We now describe a number of application examples...

Mutual Fund Sales...

Suppose a mutual fund sales person has a 50% (perhaps too high, but we will revisit this) chance of closing a sale on each call she makes. Suppose further that she made four calls in the last hour.

Consider “closing a sale” a **success** and “not closing a sale” a **failure**. Then, we will study the variables:

X = total number of successes

Y = number of successes *before* first failure

An interesting question is: How would the distribution of Y vary for *different* values of X ? This motivates the concept of a **conditional distribution**.

Conditional Probability Distribution

Formally, let X and Y be two random variables. Then, the **conditional** probability distribution of Y given $X = x$ is defined by:

$$\begin{aligned} P(y | x) &\equiv P(Y = y | X = x) \\ &= \frac{P(Y = y \text{ and } X = x)}{P(X = x)}, \end{aligned} \quad (8)$$

for all values of y .

Given $X = x$, we can also calculate the conditional expected value of Y via:

$$E(Y | X = x) = \sum_{\text{all } y} y P(y | x). \quad (9)$$

These concepts are important, particularly in regression analysis.

Details are given in **C7-01-Fund_Sales.xls**...

Lottery...

The concept of the expected value can be generalized. In many applications, we are interested in a function of a random variable. Let X be a random variable, and let $h(X)$ be a function of X ; then, the expected value of $h(X)$, written as $E(h(X))$, is defined by:

$$E(h(X)) \equiv \sum_{\text{all } x} h(x) P(x). \quad (10)$$

In a lottery where the buyer of a ticket picks 6 numbers out of 50, X can be the number of matches out of the picked numbers and the actual payoff is a function of X . We will study the expected payoff and the **risk** (standard deviation) involved in buying a lottery ticket.

Details are given in [C7-02-Lottery.xls](#)...

Managing Investments...

Managing risk is an important part of life. This is particularly true when we are assessing the desirability of an investment portfolio. It is necessary not only to look at the expected return but also to look at the risk. In this application, we will study how to reduce the risk of a portfolio.

Consider the following two investments:

	Investment 1	Investment 2
Mean Rate of Return	0.06	0.08
Standard Deviation	0.02	0.03

Which would you choose?

There is no simple answer to this question. The choice depends on your attitude toward risk.

Some people are risk averse (that is, they try to minimize their potential losses), but at the same time they may limit their potential gains. Such a person would probably pick Investment 1, since by Chebysheff's inequality, there is at least a 88.9% chance of getting a positive return on the investment (i.e., $0.06 \text{ plus/minus } 3 \cdot 0.02$).

On the other hand, people who are *not* risk averse might pick Investment 2, for although they might lose as much as 1% ($0.08 - 3 \cdot 0.03$), they might gain as much as 17% ($0.08 + 3 \cdot 0.03$).

Now, consider the following two scenarios:

Scenario 1:

	Investment 1	Investment 2
Mean Rate of Return	0.06	0.08
Standard Deviation	0.02	0.02

Scenario 2:

	Investment 1	Investment 2
Mean Rate of Return	0.08	0.08
Standard Deviation	0.02	0.03

Since the risks are the same in Scenario 1, one should pick the investment with a higher return.

In Scenario 2, the returns are the same. A risk averse person would pick the investment with a lower risk. Of course, not everyone is risk averse.

Diversification

An important idea is that one should not just invest in *one* vehicle. That is, buy a broad spectrum of stocks, bonds, real estates, money market certificates, etc. Such a strategy is called *diversifying* a portfolio.

The point here is that if you invest in, say, stocks and bonds simultaneously, then it is unlikely for them both to go down at the same time. When stock prices are increasing, bonds are usually decreasing, and vice versa. In our statistical language, this means that they are negatively correlated.

An analysis of a portfolio of 3 investments is given in [C7-03-Portfolio.xls](#)...

Odds and Subjective Probability...

In the lottery example, we found that for every dollar invested, we expected a return of only 40 cents. In other words we lost 60 cents for every dollar invested. Since one usually does not think of the lottery as a serious investment but more as a means of entertainment, this is fine. However, for actual investments we expect to get a net positive return.

In order to establish a baseline for any wager, investment, or even an insurance premium (which is a form of wager), we will study the concept of a **fair bet** in this application.

Details are given in [C7-04-Odds.xls](#)...

Decision Making under Uncertainty...

Many of the concepts we have introduced can be used effectively in analyzing decision problems that involve uncertainty.

The basic features of such problems are:

- We need to make a choice from a set of possible alternatives. Each alternative may involve a *sequence* of **actions**.
- The consequences of our actions, usually given in the form of a **payoff table**, may depend on possible **states of nature**, which are governed by a probability distribution (possibly subjective).
- The *true* state of nature is not known at the time of decision.
- Our objective is to maximize the expected payoff and/or to minimize risk.
- We could acquire additional information regarding the true state of nature at a cost.

Example: Investment Decision

An individual has \$1 million dollars and wishes to make a one-year investment.

Suppose his/her possible actions are:

a_1 : buy a guaranteed income certificate paying 10%

a_2 : buy bond with a coupon value of 8%

a_3 : buy a well-diversified portfolio of stocks

Return on investment in the diversified portfolio depends on the behavior of the interest rate next year. Suppose there are three possible states of nature:

s_1 : interest rate increases

s_2 : interest rate stays the same

s_3 : interest rate decreases

Suppose further that the subjective probabilities for these states are 0.2, 0.5, and 0.3, respectively.

Based on historical data, the payoff table is:

States of Nature	Actions		
	a_1	a_2	a_3
s_1	100,000	-50,000	150,000
s_2	100,000	80,000	90,000
s_3	100,000	180,000	40,000

Which action should he/she take? The expected payoffs for the actions are:

$$a_1: 0.2 \cdot 100,000 + 0.5 \cdot 100,000 + 0.3 \cdot 100,000 = 100,000$$

$$a_2: 0.2 \cdot (-50,000) + 0.5 \cdot 80,000 + 0.3 \cdot 180,000 = 84,000$$

$$a_3: 0.2 \cdot 150,000 + 0.5 \cdot 90,000 + 0.3 \cdot 40,000 = 87,000$$

Hence, if one wishes to maximize expected payoff, then action a_1 should be taken.

An *equivalent* concept is to minimize expected **opportunity loss** (EOL). Consider any given state. For each possible action, the opportunity loss is defined as the difference between what the payoff could have been had the *best* action been taken and the payoff for that particular action. Thus,

States of Nature	Actions		
	a_1	a_2	a_3
s_1	50,000	200,000	0
s_2	0	20,000	10,000
s_3	80,000	0	140,000
EOL:	34,000	50,000	47,000

Indeed, a_1 is again optimal.

For more complicated problems, a **decision tree** can be used. A detailed example is given in [C23-01-Decision.xls](#)...