Chapter 4 - Probability
What is $P(x = A_i)$?

Of the complete set of events in A (from A_1 to A_n) $P(x = A_i)$ is the probability that x (the variable of interest) is A_i.

Probability value ranges from 0 (x is never A_i) to 1.0 (x is A_i with complete certainty).
More fundamental properties of $P(x = A_i)$

1. $\sum_{i=1}^{n} P(x = A_i) = 1 \quad \rightarrow$ the sum of all of the probabilities equals 1 (discrete, countable case).
2. $\int_a^b P(x = A)dA = 1 \quad \rightarrow$ the sum of all of the probabilities equals 1 (continuous case).

Three ways to obtain probability values

1. One thinks of the setting as having a set of countable events. The probability of event i occurring is the proportion of event i in the set of countable events.
2. Relative frequencies from data distributions.
3. Probability distribution functions y=f(x) where y is the probability.

Determining probabilities from a set of countable events
Definitions

1. *Event* - a potential outcome from a set of outcomes.
2. *Mutually exclusive events* - describes the relationship of the outcomes in the set of outcomes. Implies that two or more outcomes cannot occur at the same time.
3. *Event space* The total set of outcomes.

In economics,
Events are referred to as States of the World and
Event space is State space.

Examples of Event spaces or state spaces

1. Two dice, set of cards, roulette wheel, urn full of balls of different colors.
2. The countable set of possible outcomes in a strategic market.
3. The countable set of possible outcomes due to uncertainty or incomplete information.

Computing Probabilities from countable events

$$P(x = A_i) = \frac{Number\ of\ events\ of\ A_i}{Maximum\ number\ of\ events}$$

Counting Rules-how does one count the number of events and the maxmimum number of events

Counting Rule #1 - consider the composition of events

1. Single stage event
2. Two Stage event
3. Multiple stage event

Counting the maximum number of events.

1. Single Stage - maximum is n where n is equal to the total number of events.
2. Two Stage - Rule → If the first stage can be accomplished in m ways and the second stage can be accomplished in n ways, there are mn possible outcomes.
3. Multiple stage - extension of two stage. If there are k steps, each with n_k steps, the total number of possible outcomes are

$$n_1 n_2 n_3 \ldots n_k$$

Counting Rule #2 Consider the number of ways we can arrange n objects, taking them r at a time

Assume three objects - red, blue, green

1. Order of each arrangement can matter - for some cases, a blue, green arrangement is different than a green, blue arrangement - **Permutation**

   Permutation

   $$P_r^n = \frac{n!}{(n-r)!}$$

2. Order does not matter - a blue, green arrangement is treated the same as a green, blue arrangement - **Combination**

Combination

$$C_r^n = \frac{n!}{r!(n-r)!}$$

Calculating probabilities using data

If a data set consists of a collection of mutually exclusive events (Categories) A_1, A_2,A_3, ...,A_n, the probability $P(x = A_i)$ is given as:

$$P(x = A_i) = \lim_{n \to \infty} \frac{Freq.\ of\ event\ A_i}{n}$$

Example #1 Tossing coins

let's look that the probability that the toss results in heads.

$$P(x = A_i(Heads)) = \lim_{tosses \to \infty} \frac{Freq.\ of\ Heads}{\#\ of\ tosses}$$

n increases as we add additional tosses to our sample. The limit condition tells us that as n increases, (e.g., as we toss the coin more), the relative frequency value will come closer to approximating the true probability of 0.50

Probability Distributions

Covered in depth in Chapters 5 and 6.

Compound Events

Compound events are form by unions or intersections of subsets of events. You can represent compound events using countable events, data distributions and probability distributions.

Example - - countable events - two coins

There are four possible arrangements of events.

HH
HT
TH
TT

Probability of any single event

$$P(x = A_i) = \frac{\text{Frequency of } A_i}{\text{Total number of events}}$$

A_i is HH, HT, TH, and TT

P(x= HH) = 0.25
P(x= HT) = 0.25
P(x= TH) = 0.25
P(x=TT) = 0.25

Suppose I have two compound events
A. Observe at least one head
B. Observe at least one tail.
Probability of x= A is equal to the probability of getting any of the following arrangements.
HH
HT
TH
Each have a probability of 0.25. The probability of x=A is the sum of these probabilities or 0.75
Probability of x=B is equal to the probability of getting any of the following arrangements.
TT
HT
TH
Each have a probability of 0.25. The probability of x=B is the sum of these probabilities or 0.75
Unions In this case, we are interested in determining the probability that the next toss is either a A or a B outcome. We examine the events that make up A and B and create a new set called $(A \cup B)$. The union set combines sets A and B and removes any duplications.
$(A \cup B)$ = P(x=HH) + P(x=HT) + P(x=TH) + P(x=TT)
= 0.25 + 0.25 + 0.25 + 0.25

= 1.0

Intersection In this case, we are interested in determining the probability that the next toss is a A outcome as well as a B outcome. This set is the intersection of A and B or $(A \cap B)$. This means that the toss must have at least one head and at least one tail. There are two events that meet this condition (HT, TH). The sum of their probabilities is 0.50

Disjoint events - a special set of compound events Disjoint sets are mutually exclusive compound sets. They are no simple events in common.

Here is an example of two disjoint sets.

C. The coins are tossed and both are head. (HH)

D. The coins are tossed and both are tail. (TT)

Probability of x=C is 0.25

Probability of x=D is 0.25

Union of C and D $(C \cup D)$ is the probability that the toss outcome is either HH or TT.

$(C \cup D)$ = P(x=C) + P(x=D) = 0.50

Intersection of C and D $(C \cap D)$ is the probability that the toss outcome of two coins has both heads and both tails. There is no simple event that matches this so the probability of the intersection of C and D or $(C \cap D) = 0$.

Complements - a special case of disjoint compound events With Complements, the simple events that are not in A are found in B.

Example of Complements

E. The coins are tossed and both are head. (HH)

F. The coins are tossed and the coins are not both heads (TT, HT, TH). (F is also called E^c or the complement of E)

Probability of x=E = 0.25

Probability of x=F = 0.75 = 1-P(x=E)

Checking the conditions for disjoint sets

P($E \cup F$) = P(x=E) + P(x=F) = 1.0

P(x=$E \cap F$) = 0

Conditional probability

Suppose you wanted to find out the probability of finding a job with a starting wage of at least H. (If we were looking at graduating econ undergrads, we might want to look at the probability of finding a jobs that pays at least $32,900). Suppose the probability is 0.50 or 50%. Now suppose we want to find out the probability that a female graduating econ undergrad finds such a job. Conditional probability can help determine this probability.

Definition - Conditional Probability

$$P(x = A|B) = \frac{P(x = A \cap B)}{P(x = B)}$$

if $P(x = B) \neq 0$

Example

	Male	Female	Total
Less than $32,900	0.28	0.22	0.50
At least $32,900	0.42	0.08	0.50
Total	0.70	0.30	

Probability of being female - P(x=F) = 0.30
Probability of earning at least $32,900 P(x=H) =0.50
$P(x = H \cap F)$ =0.08

$$P(x = H|F) = \frac{P(x = H \cap F)}{P(x = F)} = \frac{0.08}{0.30} = 0.267$$

Probability of being female - P(x=F) = 0.30
Probability of earning less than $32,900 P(x=L) =0.50
$P(x = L \cap F)$ =0.22

$$P(x = L|F) = \frac{P(x = L \cap F)}{P(x = F)} = \frac{0.22}{0.30} = 0.733 = (1 - P(x = H|F))$$

Conditional probability outcomes tells you if the two sets of events are independent from one another

Independent - the knowledge of conditional event does not change the probability of the event.

$$P(x = A|B) = P(x = A)$$

Suppose the following assumptions held.

	Male	Female	Total
Less than \$32,900	0.35	0.15	0.50
At least \$32,900	0.35	0.15	0.50
Total	0.70	0.30	

Probability of being female - P(x=F) = 0.30
Probability of earning at least \$32,900 P(x=W) =0.50
$P(x = F \cap W)$ =0.15

$$P(x = W|F) = \frac{P(x = F \cap W)}{P(x = F)} = \frac{0.15}{0.30} = 0.50 = P(x = W)$$

Conclusion: For this select population, gender and wage rate are independent from one another.

Two additional rules

a) Additive Rule

$$P(x = A \cup B) = P(x = A) + P(x = B) - P(x = A \cap B)$$

b) Multiplicative Rule

$$P(x = A \cap B) = P(x = A)P(x = B|A)$$

Examples

	Male	Female	Total
Less than \$32,900 (L)	0.28	0.22	0.50
At least \$32,900 (H)	0.42	0.08	0.50
Total	0.70	0.30	

Probability that the person is female or earning less than \$32,900. P(x=$F \cup L$).

$$P(x = F \cup L) = P(x = F) + P(x = L) - P(x = F \cap L) = 0.30 + 0.50 - 0.22 = 0$$

Probability that the person is female and earning at least \$32,900. P($x = F \cap H$).

$$P(x = F \cap H) = P(x = F)P(x = W|F) = 0.30 * 0.267 = 0.08$$

Discrete Random Variables and their probability distributions Previously we calculate relative frequencies of categories of variables. If a certain condition hold for these frequencies, we can refer to these values as probabilities. This introduces two new statistical tools - formulas for calculating mean and variance from relative frequency tables with discrete categories. The necessary condition to use these tools is that the variable is a random variable.

Random Variable

A variable is a random variable if the value it assumes is a chance event.

Important

A variable that is a random variable has a distribution function or a probability distribution function.

Discrete Distributions

A distribution is discrete if the set of values are quantitative and countable. Here are some examples.

1. Age of the labor force.
2. Number of years of formal school - labor force.
3. Number of countries that Country X trades with.

Examples of distributions that are not discrete

1. Values represents as a rate, e.g., Change in GDP, unemployment rate, interest rate.
2. Income

Properties of Discrete Probability Distribution

1. $0 \leq p(x_i) \leq 1$
2. $\sum_{i=1}^{N} p(x_i) = 1$

Mean and Standard Deviation- discrete random variable

If we know each value of X $(x_1, x_2, \dots x_N)$ and each probability value $(p(x_1), p(x_2), \dots p(x_N))$, we can compute its mean and standard deviation.

$$\mu = \sum_{i=1}^{N} x_i p(x_i)$$

$$\sigma^2 = \sum_{i=1}^{N} (x_i - \mu)^2 p(x_i)$$

Examples