GOV 2000 / E-2000 Section 3[1]

Solé Prillaman

Harvard University

September 19, 2013

[1] These notes and accompanying code draw on the notes from TF's from previous years.

LOGISTICS

Problem Set 2- Due by 1pm on Tuesday to course dropbox.

Problem Set 1 Corrections- Due by 1pm on Tuesday to course dropbox.

Reading Quiz- Due by 9pm on Sunday on course website.

Logistics

Problem Set 2- Due by 1pm on Tuesday to course dropbox.

Problem Set 1 Corrections- Due by 1pm on Tuesday to course dropbox.

Reading Quiz- Due by 9pm on Sunday on course website.

LaTeX Tutorial- Next Friday, September 27, at 4pm. Location TBD.

Outline

Quick R Code

Contents of your lm object:

```
## load some data
load(file="Leinhardt.RData")

## run a model
my.lm <- lm(linfant ~ lincome, data=Leinhardt)

## our lm object has many components we may want to extract
names(my.lm)

## Two ways to get the coefficients:
my.coefs <- my.lm$coefficients
my.coefs <- coefficients(my.lm)

## Two ways to get the residuals:
my.resids <- my.lm$residuals
my.resids <- residuals(my.lm)

## One way to get the fitted values:
my.fvs <- my.lm$fitted.values
```

Outline

Why Add More Variables?

1. Improve fit of model.

2. Theory calls for it.

Review: Additive Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z}$$

[2] Note that this is not a causal effect, just an expected relationship. We will discuss this more later in the semester. See the interpretation slide for a more precise definition.

Review: Additive Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z}$$

What is the effect of X on Y?[2]

[2] Note that this is not a causal effect, just an expected relationship. We will discuss this more later in the semester. See the interpretation slide for a more precise definition.

Review: Additive Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z}$$

What is the effect of X on Y?[2] β_1

[2] Note that this is not a causal effect, just an expected relationship. We will discuss this more later in the semester. See the interpretation slide for a more precise definition.

Review: Additive Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z}$$

What is the effect of X on Y?[2] β_1

What is the intercept when plotting Y on X?

[2] Note that this is not a causal effect, just an expected relationship. We will discuss this more later in the semester. See the interpretation slide for a more precise definition.

Review: Additive Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z}$$

What is the effect of X on Y?[2] β_1

What is the intercept when plotting Y on X? $\beta_0 + \beta_2 Z$

[2] Note that this is not a causal effect, just an expected relationship. We will discuss this more later in the semester. See the interpretation slide for a more precise definition.

Additive Multiple Regression with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z$.

What is the relationship between X and Y?
(Suppose Z can be any positive integer)

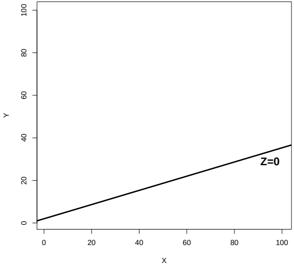
Z	Intercept $\beta_0 + \beta_2 Z$	Slope β_1
0		
1		
2		
3		
$\vdots$	$\vdots$	$\vdots$
1253		

Additive Multiple Regression with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z$.

What is the relationship between X and Y?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope β_1
0	2	$\frac{1}{3}$
1		
2		
3		
$\vdots$	$\vdots$	$\vdots$
1253		

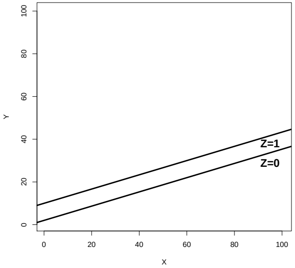


Additive Multiple Regression with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z$.

What is the relationship between X and Y?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope β_1
0	2	$\frac{1}{3}$
1	2 + 8	$\frac{1}{3}$
2		
3		
$\vdots$	$\vdots$	$\vdots$
1253		

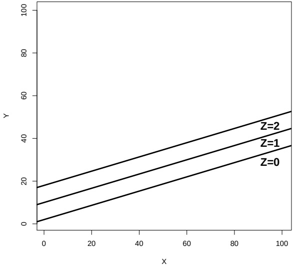


Additive Multiple Regression with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z$.

What is the relationship between X and Y?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope β_1
0	2	$\frac{1}{3}$
1	$2+8$	$\frac{1}{3}$
2	$2+8\cdot 2$	$\frac{1}{3}$
3		
$\vdots$	$\vdots$	$\vdots$
1253		

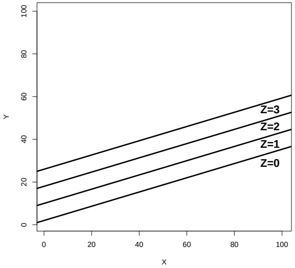


Additive Multiple Regression with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z$.

What is the relationship between X and Y?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope β_1
0	2	$\frac{1}{3}$
1	$2 + 8$	$\frac{1}{3}$
2	$2 + 8 \cdot 2$	$\frac{1}{3}$
3	$2 + 8 \cdot 3$	$\frac{1}{3}$
$\vdots$	$\vdots$	$\vdots$
1253		

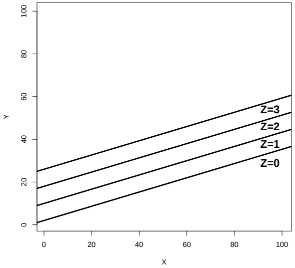


Additive Multiple Regression with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z$.

What is the relationship between X and Y?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope β_1
0	2	$\frac{1}{3}$
1	$2 + 8$	$\frac{1}{3}$
2	$2 + 8 \cdot 2$	$\frac{1}{3}$
3	$2 + 8 \cdot 3$	$\frac{1}{3}$
$\vdots$	$\vdots$	$\vdots$
1253	$2 + 8 \cdot 1253$	$\frac{1}{3}$



Interpretation

How do we interpret β_1 in the additive multiple regression model?

Interpretation

How do we interpret β_1 in the additive multiple regression model?

The expected change in Y associated with a one-unit increase in X, holding constant the value of Z.

Review: Interaction Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \cdot Z}$$

Review: Interaction Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \cdot Z}$$

What is the conditional effect of X on Y?

Review: Interaction Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \cdot Z}$$

What is the conditional effect of X on Y? $\beta_1 + \beta_3 Z$

Review: Interaction Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \cdot Z}$$

What is the conditional effect of X on Y? $\beta_1 + \beta_3 Z$

$$\begin{aligned}
Y &= \beta_0 + \beta_1 X + \beta_3 X \cdot Z + \beta_2 Z \\
Y &= \beta_0 + (\beta_1 + \beta_3 Z)X + \beta_2 Z \\
Y &= (\beta_0 + \beta_2 Z) + (\beta_1 + \beta_3 Z)X
\end{aligned}$$

Review: Interaction Models

$$\boxed{\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X \cdot Z}$$

What is the conditional effect of X on Y? $\beta_1 + \beta_3 Z$

$$\begin{aligned}
Y &= \beta_0 + \beta_1 X + \beta_3 X \cdot Z + \beta_2 Z \\
Y &= \beta_0 + (\beta_1 + \beta_3 Z)X + \beta_2 Z \\
Y &= (\beta_0 + \beta_2 Z) + (\beta_1 + \beta_3 Z)X
\end{aligned}$$

or...

$$\frac{\partial Y}{\partial X} = \beta_1 + \beta_3 Z$$

Quiz Break!

What is the conditional intercept when plotting Y on X?

What is the conditional effect of Z on Y?

Quiz Break!

What is the conditional intercept when plotting Y on X?

$\beta_0 + \beta_2 Z$

What is the conditional effect of Z on Y?

Quiz Break!

What is the conditional intercept when plotting Y on X?

$\beta_0 + \beta_2 Z$

What is the conditional effect of Z on Y? $\beta_2 + \beta_3 X$

Interactions with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z + \frac{1}{2}X \cdot Z$.

What is the relationship between X and Y conditional on Z?
(Suppose Z can be any positive integer)

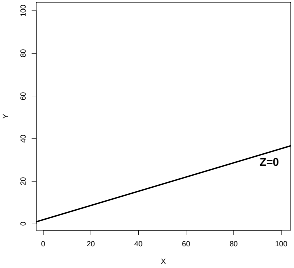
Z	Intercept $\beta_0 + \beta_2 Z$	Slope $\beta_1 + \beta_3 Z$
0		
1		
2		
3		
$\vdots$	$\vdots$	$\vdots$
1253		

Interactions with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z + \frac{1}{2}X \cdot Z$.

What is the relationship between X and Y conditional on Z?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope $\beta_1 + \beta_3 Z$
0	2	$\frac{1}{3}$
1		
2		
3		
$\vdots$	$\vdots$	$\vdots$
1253		

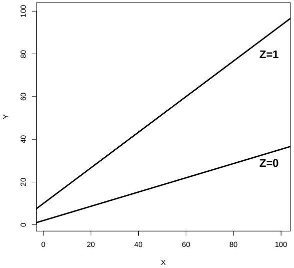


INTERACTIONS WITH TWO CONTINUOUS VARIABLES

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z + \frac{1}{2}X \cdot Z$.

What is the relationship between X and Y conditional on Z?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope $\beta_1 + \beta_3 Z$
0	2	$\frac{1}{3}$
1	2 + 8	$\frac{1}{3} + \frac{1}{2}$
2		
3		
$\vdots$	$\vdots$	$\vdots$
1253		

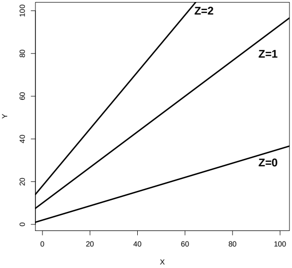


Interactions with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z + \frac{1}{2}X \cdot Z$.

What is the relationship between X and Y conditional on Z? (Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope $\beta_1 + \beta_3 Z$
0	2	$\frac{1}{3}$
1	$2+8$	$\frac{1}{3} + \frac{1}{2}$
2	$2+8\cdot 2$	$\frac{1}{3} + \frac{1}{2} \cdot 2$
3		
$\vdots$	$\vdots$	$\vdots$
1253		

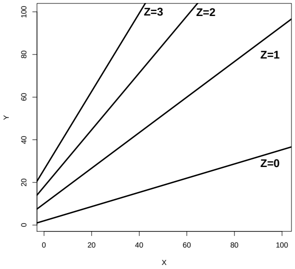


Interactions with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z + \frac{1}{2}X \cdot Z$.

What is the relationship between X and Y conditional on Z?
(Suppose Z can be any positive integer)

Z	Intercept $\beta_0 + \beta_2 Z$	Slope $\beta_1 + \beta_3 Z$
0	2	$\frac{1}{3}$
1	$2+8$	$\frac{1}{3} + \frac{1}{2}$
2	$2+8\cdot 2$	$\frac{1}{3} + \frac{1}{2} \cdot 2$
3	$2+8\cdot 3$	$\frac{1}{3} + \frac{1}{2} \cdot 3$
$\vdots$	$\vdots$	$\vdots$
1253		

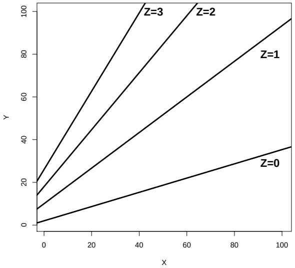


Interactions with Two Continuous Variables

We estimate that $\hat{Y} = 2 + \frac{1}{3}X + 8Z + \frac{1}{2}X \cdot Z$.

What is the relationship between X and Y conditional on Z?
(Suppose Z can be any positive integer)

Z	Intercept	Slope
	$\beta_0 + \beta_2 Z$	$\beta_1 + \beta_3 Z$
0	2	$\frac{1}{3}$
1	$2+8$	$\frac{1}{3} + \frac{1}{2}$
2	$2+8\cdot 2$	$\frac{1}{3} + \frac{1}{2} \cdot 2$
3	$2+8\cdot 3$	$\frac{1}{3} + \frac{1}{2} \cdot 3$
$\vdots$	$\vdots$	$\vdots$
1253	$2 + \frac{4}{3} \cdot 1253$	$\frac{1}{2} + 8 \cdot 1253$



INTERPRETATION

How do we interpret $\beta_1 + \beta_3 z$ in the interactive multiple regression model?

INTERPRETATION

How do we interpret $\beta_1 + \beta_3 z$ in the interactive multiple regression model?

The expected change in Y associated with a one-unit increase in X, conditional on Z being z.

When Interaction?

When the hypotheses we are testing are of the form:[3]

> *1. The effect of X on Y is increasing(decreasing) in Z. An increase in X is associated with a larger(smaller) increase in Y when condition Z is met, than when condition Z is absent.*

Example?

[3] See Brambor (2006) for more.

WHEN INTERACTION?

When the hypotheses we are testing are of the form:[3]

1. The effect of X on Y is increasing(decreasing) in Z. An increase in X is associated with a larger(smaller) increase in Y when condition Z is met, than when condition Z is absent.

Example?

More education is associated with higher wages when "ability" is low. More education is associated with even higher wages when "ability" is high.

[3] See Brambor (2006) for more.

When Interaction?

When the hypotheses we are testing are of the form:[3]

> *1. The effect of X on Y is increasing(decreasing) in Z. An increase in X is associated with a larger(smaller) increase in Y when condition Z is met, than when condition Z is absent.*

Example?

More education is associated with higher wages when "ability" is low. More education is associated with even higher wages when "ability" is high.

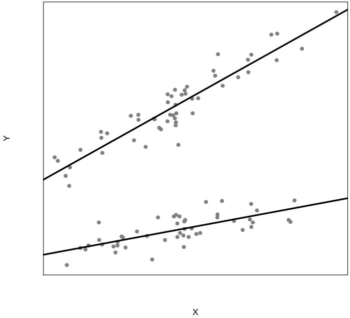
$$Wages = \beta_0 + \beta_1 education + \beta_2 ability + \beta_3 education \cdot ability$$

[3] See Brambor (2006) for more.

When Interaction?

When the hypotheses we are testing are of the form:

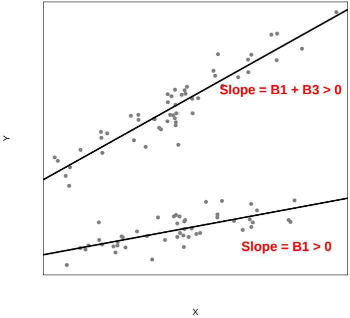
1. The effect of X on Y is increasing(decreasing) in Z. An increase in X is associated with a larger(smaller) increase in Y when condition Z is met, than when condition Z is absent.



When Interaction?

When the hypotheses we are testing are of the form:

1. The effect of X on Y is increasing(decreasing) in Z. An increase in X is associated with a larger(smaller) increase in Y when condition Z is met, than when condition Z is absent.



When Interaction?

When the hypotheses we are testing are of the form:

> *2. An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.*

Example?

[4] See Boix (1999).

When Interaction?

When the hypotheses we are testing are of the form:

> *2. An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.*

Example?

High ethnic or religious fragmentation encourages the adoption of PR systems in **small** countries, but not in large countries.[4]

[4] See Boix (1999).

WHEN INTERACTION?

When the hypotheses we are testing are of the form:

> *2. An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.*

Example?

High ethnic or religious fragmentation encourages the adoption of PR systems in **small** countries, but not in large countries.[4]

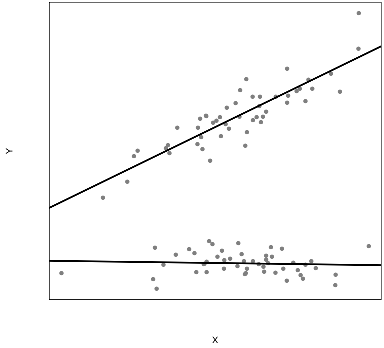
$$PR = \beta_0 + \beta_1 fragmentation + \beta_2 countrysize + \beta_3 fragmentation \cdot countrysize$$

[4] See Boix (1999).

WHEN INTERACTION?

When the hypotheses we are testing are of the form:

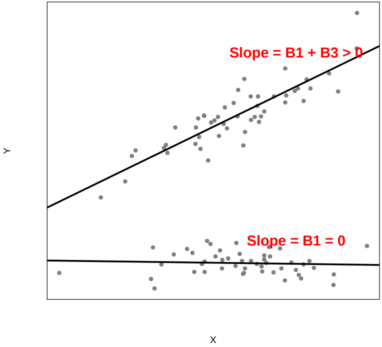
> *2. An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.*



When Interaction?

When the hypotheses we are testing are of the form:

2. An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.



WHEN INTERACTION?

When the hypotheses we are testing are of the form:

> *3. An increase in X is associated with an increase in Y when condition Z is present, but a decrease in Y when condition Z is absent.*

Example?

When Interaction?

When the hypotheses we are testing are of the form:

> *3. An increase in X is associated with an increase in Y when condition Z is present, but a decrease in Y when condition Z is absent.*

Example?

Aspirin is expected to make you feel better when you have not been drinking and is expected to make you feel worse when you have been drinking.

When Interaction?

When the hypotheses we are testing are of the form:

> *3. An increase in X is associated with an increase in Y when condition Z is present, but a decrease in Y when condition Z is absent.*

Example?

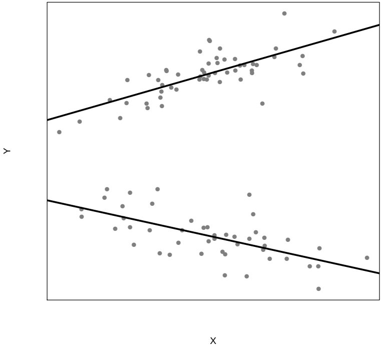
Aspirin is expected to make you feel better when you have not been drinking and is expected to make you feel worse when you have been drinking.

$$feel = \beta_0 + \beta_1 Aspirin + \beta_2 drinking + \beta_3 Aspirin \cdot drinking$$

WHEN INTERACTION?

When the hypotheses we are testing are of the form:

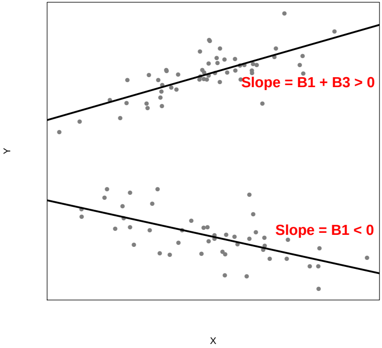
> *3. An increase in X is associated with an increase in Y when condition Z is present, but a decrease in Y when condition Z is absent.*



When Interaction?

When the hypotheses we are testing are of the form:

> *3. An increase in X is associated with an increase in Y when condition Z is present, but a decrease in Y when condition Z is absent.*



Questions?

Outline

PROBABILITY

Probability is a formal model of uncertainty.

Suppose I roll a die.

What is the probability we roll a 2?

Probability

Probability is a formal model of uncertainty.

Suppose I roll a die.

What is the probability we roll a 2?
$P(X = 2) = \frac{1}{6}$

PROBABILITY

Probability is a formal model of uncertainty.

Suppose I roll a die.

What is the probability we roll something less than 4?

Probability

Probability is a formal model of uncertainty.

Suppose I roll a die.

What is the probability we roll something less than 4?
$P(X < 4) = P(X = 1) + P(X = 2) + P(X + 3)$

Probability

Probability is a formal model of uncertainty.

Suppose I roll a die.

What is the probability we roll something less than 4?
$P(X < 4) = P(X = 1) + P(X = 2) + P(X + 3) = \frac{3}{6}$

Notation and Definitions

Sample Space (S): The set of all possible outcomes from some process.

Notation and Definitions

Sample Space (S): The set of all possible outcomes from some process.

Event: Any subset of the sample space.

Notation and Definitions

Sample Space (S): The set of all possible outcomes from some process.

Event: Any subset of the sample space.

Union: The union of two events A and B, $A \cup B$, is the set containing all elements of A **or** B.

Notation and Definitions

Sample Space (S): The set of all possible outcomes from some process.

Event: Any subset of the sample space.

Union: The union of two events A and B, $A \cup B$, is the set containing all elements of A **or** B.

Intersection: The intersection of two events A and B, $A \cap B$, is the set containing all elements in both A **and** B.

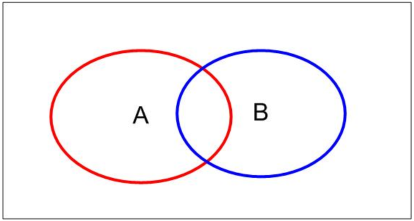
Notation and Definitions

Sample Space (S): The set of all possible outcomes from some process.

Event: Any subset of the sample space.

Union: The union of two events A and B, $A \cup B$, is the set containing all elements of A **or** B.

Intersection: The intersection of two events A and B, $A \cap B$, is the set containing all elements in both A **and** B.



Axioms of Probability

1. For any event A, $P(A) \geq 0$
2. $P(S) = 1$ where S is the sample space
3. For any sequence of disjoint events (A and B), $P(A \cup B) = P(A) + P(B)$

Probability Example

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Let's define C to be the event of having a car phone and S to be the event of receiving a speeding violation.

Marginal Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Marginal Probability: The probability of an event without including any additional information.

Marginal Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Marginal Probability: The probability of an event without including any additional information. Look at the margins!

Marginal Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Marginal Probability: The probability of an event without including any additional information.Look at the margins!

What is the probability of having a car phone?

Marginal Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Marginal Probability: The probability of an event without including any additional information. Look at the margins!

What is the probability of having a car phone? $P(C) = \frac{305}{755}$

Joint Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Joint Probability: The probability of some event A and another event B both occurring (P(A∩B) or P(A,B)).

Joint Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Joint Probability: The probability of some event A and another event B both occurring (P(A∩B) or P(A,B)).Look inside the table.

Joint Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Joint Probability: The probability of some event A and another event B both occurring ($P(A \cap B)$ or $P(A,B)$). Look inside the table.

What is the probability of getting a speeding violation and having a car phone?

Joint Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Joint Probability: The probability of some event A and another event B both occurring (P(A∩B) or P(A,B)).Look inside the table.

What is the probability of getting a speeding violation and having a car phone? $P(S \cap C) = \frac{25}{755}$

Conditional Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Conditional Probability: The probability of some event A, given that another event B has occurred ($P(A|B)$).

Conditional Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Conditional Probability: The probability of some event A, given that another event B has occurred (P(A|B)). Condition on the row/column and then calculate the probability.

Conditional Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Conditional Probability: The probability of some event A, given that another event B has occurred (P(A|B)).Condition on the row/column and then calculate the probability.

What is the probability of getting a speeding violation given that you have a car phone?

Conditional Probability

	Speeding Violation	No Speeding Violation	Total
Car Phone	25	280	305
No Car Phone	45	405	450
Total	70	685	755

Source: `http://cnx.org/content/m16835/1.12/`

Conditional Probability: The probability of some event A, given that another event B has occurred ($P(A|B)$). Condition on the row/column and then calculate the probability.

What is the probability of getting a speeding violation given that you have a car phone? $P(S|C) = \frac{25}{305}$

Conditional Probability

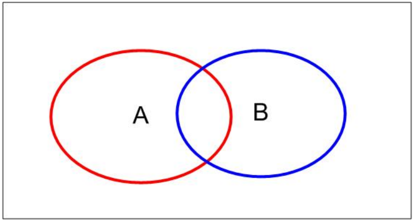
What if we don't have such a nice table?

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

Conditional Probability

What if we don't have such a nice table?

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



CONDITIONAL PROBABILITY

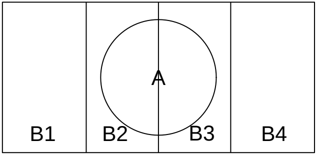
Note that if:

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} \\ P(B|A)P(A) &= P(A \cap B) \end{aligned}$$

Law of Total Probability

Law of Total Probability: Let S be the sample space and let the disjoint k events $B_1, ..., B_k$ partition S such that $P(B_1 \cup ... \cup B_k) = P(S) = 1$. If A is some other event in S, then

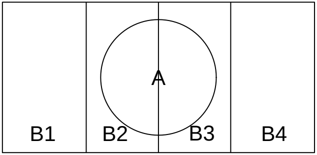
$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$$



Law of Total Probability

Law of Total Probability: Let S be the sample space and let the disjoint k events $B_1, \ldots, B_k$ partition S such that P($B_1 \cup \ldots \cup B_k$) $= P(S) = 1$. If A is some other event in S, then

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(B_i) P(A|B_i)$$

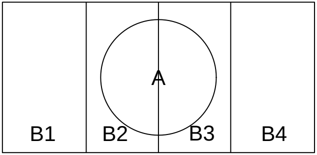


What is the P(A) in this figure?

Law of Total Probability

Law of Total Probability: Let S be the sample space and let the disjoint k events $B_1, ..., B_k$ partition S such that P$(B_1 \cup ... \cup B_k) = P(S) = 1$. If A is some other event in S, then

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(B_i)P(A|B_i)$$



What is the P(A) in this figure?

$P(A) =$
$P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) + P(B_4)P(A|B_4)$

Bayes' Rule

$$
\begin{aligned}
P(B|A) &= \frac{P(A \cap B)}{P(A)} && \text{Definition of Conditional Prob.} \\
&= \frac{P(B)P(A|B)}{P(A)} && \text{Rearranged Conditional Prob.} \\
&= \frac{P(B)P(A|B)}{P(B)P(A|B) + P(B^C)P(A|B^C)} && \text{Law of Total Prob.}
\end{aligned}
$$

Quiz Break!

Write a function in R that takes in $P(A)$, $P(B)$, and $P(A|B)$ and outputs $P(B|A)$.

OUTLINE

Discrete Random Variables

Random Variable: a real-valued function mapping the sample space S to a finite number of distinct values.

Discrete Random Variables

Random Variable: a real-valued function mapping the sample space S to a finite number of distinct values.

Probability Mass Function: assigns probabilities to a countable number of distinct values; the probability that some random variable X is equal to some value x ($P(X = x)$).

Discrete Random Variables

Random Variable: a real-valued function mapping the sample space S to a finite number of distinct values.

Probability Mass Function: assigns probabilities to a countable number of distinct values; the probability that some random variable X is equal to some value x ($P(X = x)$).

Cumulative Density Function: the probability that some random variable X is less than or equal to some value x ($P(X \leq x)$).

Discrete Distributions

Imagine that you roll a die continuously and note the outcome of the roll.

Discrete Distributions

Imagine that you roll a die continuously and note the outcome of the roll.

What is the random variable? (Call it X)

Discrete Distributions

Imagine that you roll a die continuously and note the outcome of the roll.

What is the random variable? (Call it X) The outcome of a roll of a die.

Discrete Distributions

Imagine that you roll a die continuously and note the outcome of the roll.

What is the random variable? (Call it X) The outcome of a roll of a die.

What is the distribution of X?

Discrete Distributions

Imagine that you roll a die continuously and note the outcome of the roll.

What is the random variable? (Call it X) The outcome of a roll of a die.

What is the distribution of X? $P(X) = \begin{cases} \frac{1}{6} & \text{if } x = 1 \\ \frac{1}{6} & \text{if } x = 2 \\ \frac{1}{6} & \text{if } x = 3 \\ \frac{1}{6} & \text{if } x = 4 \\ \frac{1}{6} & \text{if } x = 5 \\ \frac{1}{6} & \text{if } x = 6 \\ 0 & \text{otherwise} \end{cases}$

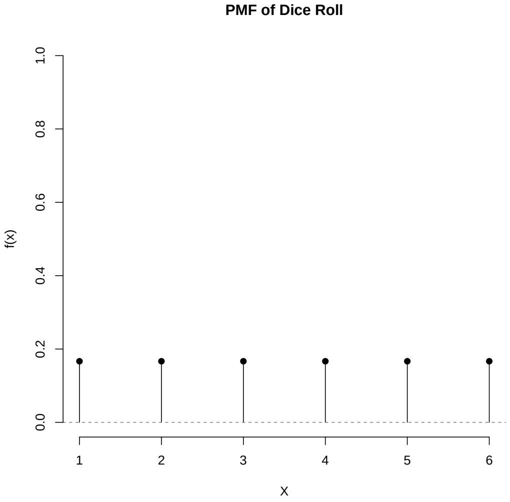
Discrete Distributions in R

How would we draw this PMF?

```
## Create an empty plot
plot(0,0,type="n",xlab="X",ylab="f(x)",
     xlim=c(1,6), ylim=c(0,1),
     main="PMF of Dice Roll")

## Add each vertical line and each point
## for each line we specify the starting point (x0,y0)
##  and end point(x1,y1)
segments(x0=1,x1=1,y0=0,y1=1/6)
points(x=1,y=1/6,pch=19)
segments(x0=2,x1=2,y0=0,y1=1/6)
points(x=2,y=1/6,pch=19)
segments(x0=3,x1=3,y0=0,y1=1/6)
points(x=3,y=1/6,pch=19)
segments(x0=4,x1=4,y0=0,y1=1/6)
points(x=4,y=1/6,pch=19)
segments(x0=5,x1=5,y0=0,y1=1/6)
points(x=5,y=1/6,pch=19)
segments(x0=6,x1=6,y0=0,y1=1/6)
points(x=6,y=1/6,pch=19)
```

Discrete Distributions in R

How would we draw the PMF?



Discrete Distributions in R

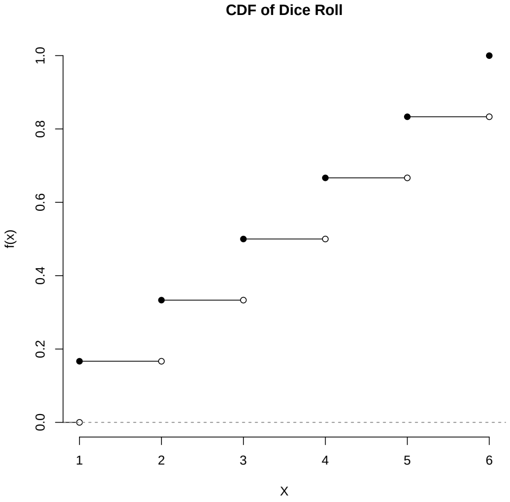
How would we draw the CDF?

```
## Create an empty plot
plot(0,0,type="n",xlab="X",ylab="f(x)",
     xlim=c(1,6), ylim=c(0,1),
     main="PMF of Dice Roll")

## First, add the horizontal line for each value of x
segments(x0=0,x1=1,y0=0,y1=0)

## Then add the open point
points(x=1,y=0,pch=21,bg="white")

## Then add the closed point for the cumulative probability
points(x=1,y=1/6,pch=19)

## Repeat for all values of x
```

See if you can finish this CDF on your own!

Discrete Distributions in R

How would we draw the CDF?



Discrete Distributions in R

What if we aren't certain about the distribution of X?

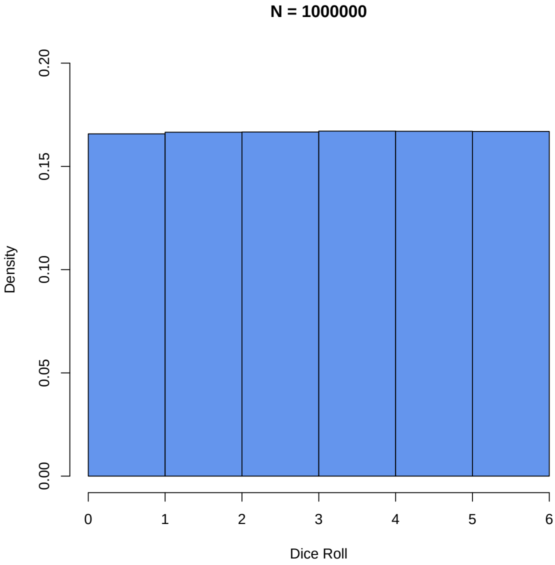
Discrete Distributions in R

What if we aren't certain about the distribution of X?

We can sample from our population or sample space and then look at the distribution of our sample.

```
## Define the sample space
S <- 1:6

## Set the sample size
n <- 1000000

## Sample from the population
## i.e. Roll the "die" repeatedly
dice.rolls <- sample(S, size=n, replace=T)

## Plot the distribution
hist(dice.rolls, freq=F)
```

Discrete Distributions in R



Continuous Random Variables

Continuous Random Variable: a real-valued function mapping the sample space S to the real line.

Continuous Random Variables

Continuous Random Variable: a real-valued function mapping the sample space S to the real line.

Probability Density Function: the function f is a probability density function of some random variable X if $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$.

CONTINUOUS RANDOM VARIABLES

Continuous Random Variable: a real-valued function mapping the sample space S to the real line.

Probability Density Function: the function f is a probability density function of some random variable X if $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$.

Cumulative Density Function: the probability that some random variable X is on the interval $(-\infty, x)$ $(P(X \leq x))$.

Continuous Distributions in R

Suppose our random variable of interest is people's salary.

CONTINUOUS DISTRIBUTIONS IN R

Suppose our random variable of interest is people's salary.

Let's assume that this random variable is normally distributed with mean salary $40,000$ and standard deviation $10,000$.

Continuous Distributions in R

Suppose our random variable of interest is people's salary.

Let's assume that this random variable is normally distributed with mean salary $40,000$ and standard deviation $10,000$.

What if we want to look at the distribution of X?

Continuous Distributions in R

Suppose our random variable of interest is people's salary.

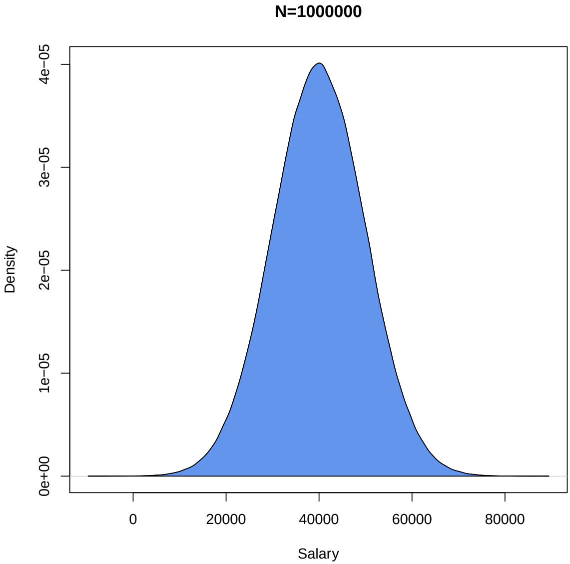
Let's assume that this random variable is normally distributed with mean salary $40,000$ and standard deviation $10,000$.

What if we want to look at the distribution of X?

Again, we can sample from our population or sample space and then look at the distribution of our sample.

```
## Sample from the population
salaries <- rnorm(n=1000000 , mean=40000 , sd=10000)

## Plot the distribution
plot(density(salaries))
```

Continuous Distributions in R



Other Cool Things in R

What is f(20000)?

```
## dnorm provides the PDF of the normal distribution
dnorm(20000, mean=40000, sd=10000)
```

Other Cool Things in R

What is f(20000)?

```
## dnorm provides the PDF of the normal distribution
dnorm(20000, mean=40000, sd=10000)
```

What is the probability that a salary is less that $20,000?

```
## pnorm provides the CDF of the normal distribution
pnorm(20000, mean=40000, sd=10000)
```

Other Cool Things in R

What is f(20000)?

```
## dnorm provides the PDF of the normal distribution
dnorm(20000, mean=40000, sd=10000)
```

What is the probability that a salary is less that $20,000?

```
## pnorm provides the CDF of the normal distribution
pnorm(20000, mean=40000, sd=10000)
```

What value of p satisfies $P(X \leq p) = .95$?

```
## qnorm provides the inverse CDF of the normal distribution
qnorm(0.95, mean=40000, sd=10000)
```

Other Cool Things in R

For a list of the distributions that have similar commands in R, go to `http://www.stat.umn.edu/geyer/old/5101/rlook.html`.

Outline

Expectation

Expectation of a Discrete Random Variable:
$E[X] = \sum_x x \cdot P(X = x)$

Conditional Expectation of a Discrete Random Variable:
$E[X|Y = y] = \sum_x x \cdot P(X = x|Y = y)$

Expectation of a Continuous Random Variable:
$E[X] = \int_x x \cdot f(x) dx$

Conditional Expectation of a Discrete Random Variable:
$E[X|Y = y] = \int_x x \cdot f(x|y) dx$

Variance

Variance of a Discrete Random Variable:
$V[X] = \sum_x (x - E[X])^2 \cdot P(X = x)$

Variance of a Continuous Random Variable:
$V[X] = \int_x (x - E[X])^2 \cdot f(x) dx$

Outline

Questions?

REFERENCES

Boix, Carles. 1999. "Setting the Rules of the Games: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review* 93:609-624.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14:63-82.