# Probably About Probability

$\underline{p} < .05$

Greg C Elvers

1

## Probability

- Inferential statistics allow us to decide if one condition of an experiment produces different results than another condition
- Inferential statistics are based on the concepts of probability
- Thus, probability is an essential aspect of statistics

2

## What Is Probability?

- Probabilities often deal with *events*
- An event is something that happens
  - E.g. Rolling a 3 on a fair die is an event
- The probability of an event is given by the ratio of how often that event occurs and how often all events occur

3

## Probability of Events

- When you role a fair, 6 sided die, each of the six faces has an equal chance of coming up
- Thus, the probability of any single face appearing is given by 1 (how often that event occurs) divided by 6 (the total number of events)

| Event | p(Event) |
|---|---|
| 1 | 1/6 |
| 2 | 1/6 |
| 3 | 1/6 |
| 4 | 1/6 |
| 5 | 1/6 |
| 6 | 1/6 |

4

## Probability of Events

- What is the probability of *not* rolling a 1?
- There are five events that are not 1
- There is a total of six events
- The probability of not rolling a 1 is 5 / 6
- $p(\bar{1}) = 5 / 6$

| Event | p(Event) |
|---|---|
| 1 | 1/6 |
| 2 | 1/6 |
| 3 | 1/6 |
| 4 | 1/6 |
| 5 | 1/6 |
| 6 | 1/6 |



## The Addition Rule

- The *addition rule* is used to determine the probability of two or more events occurring
  - E.g., what is the probability that an odd number will appear on the die?
- For *mutually exclusive* events, the addition rule is:
  p(A or B) = p(A) + p(B)
- Two events are mutually exclusive when both cannot occur at the same time



## The Addition Rule

- An “odd” event occurs whenever the die comes up 1, 3, or 5
- These events are mutually exclusive
  - E.g., if it comes up 3, it cannot also be 1 or 5
- p(1 or 3 or 5) = p(1) + p(3) + p(5) = 1/6 + 1/6 + 1/6 = .5

| Event | p(Event) |
|---|---|
| 1 | 1/6 |
| 2 | 1/6 |
| 3 | 1/6 |
| 4 | 1/6 |
| 5 | 1/6 |
| 6 | 1/6 |



## The Addition Rule

- When events are not mutually exclusive, a different addition rule must be used
  - When events are not mutually exclusive, one or more of the events can occur at the same time
- Are the events “liking cats” and “liking stats” mutually exclusive?

| | Like Cats | Dislike Cats |
|---|---|---|
| Like Stats | 4 | 3 |
| Dislike Stats | 2 | 5 |

# The Addition Rule

- The two events, "liking stats" and "liking cats" are not mutually exclusive
- It is possible for a person to like statistics and either like or dislike cats

| | Like Cats | Dislike Cats |
|---|---|---|
| Like Stats | 4 | 3 |
| Dislike Stats | 2 | 5 |



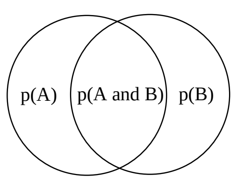
# The Addition Rule

- When events are not mutually exclusive, the addition rule is given by:
- p(A or B) = p(A) + p(B) - p(A and B)
- p(A and B) is the probability that both event A and event B occur simultaneously
- This formula can always be used as the addition rule because p(A and B) equals zero when the events are mutually exclusive



# The Addition Rule

- Why do we subtract p(A and B)?
- When the events are not mutually exclusive, some events that are A, are also B
- Those events are counted twice in p(A) + p(B)
- p(A and B) removes the second counting of the events that are both A and B






# The Multiplication Rule

- To determine the probability of two (or more) independent events occurring simultaneously, one uses the multiplication rule
- p(A and B) = p(A) X p(B)
- Note: This formula can be used to solve the previous addition rule, *but only if the events are independent*

## Independent Events

- Two events are said to be *independent* if the occurrence of one event in no way influences the occurrence of the other event
- That is, knowing something about whether one event has occurred tells you nothing about whether the other event has occurred
  - E.g., flipping a coin twice
  - E.g., being struck by lightning and having green eyes

13

## The Multiplication Rule

- What is the probability of rolling two dice and having both show 6?
- p(6 and 6) = p(6) X p(6) = 1/6 X 1/6 = 1/36

| Event | p(Event) |
|---|---|
| 1 | 1/6 |
| 2 | 1/6 |
| 3 | 1/6 |
| 4 | 1/6 |
| 5 | 1/6 |
| 6 | 1/6 |

14

## Some Probability Problems

- What is the probability of selecting a King in a single draw from a standard deck of cards? Solution
- What is the probability of selecting a face card (Jack, Queen, or King) in a single draw from a deck of cards? Solution
- Four cards are dealt from a deck with replacement. What is the probability that all four cards are aces? Solution

15

## Joint and Marginal Probabilities

- A *joint probability* is the probability of two (or more) events happening together
  - E.g. The probability that a person likes statistics and likes cats
- A *marginal probability* is the probability of just one of those events
  - E.g. probability of liking statistics

16

## Joint and Marginal Probabilities

| | Likes Cats (B) | Does Not Like Cats (not B) | Marginal Probability |
|---|---|---|---|
| Likes Stat (A) | 3<br>.167 | 1<br>.056 | 4<br>.222 |
| Does Not Like Stats (not A) | 10<br>.556 | 4<br>.222 | 14<br>.778 |
| Marginal Probability | 13<br>.722 | 5<br>.278 | 18 |



## Conditional Probability

- If two events are *not* independent of each other, then knowing whether one event occurred changes the probability that the other event might occur
  - E.g., knowing that a person is an introvert decreases the probability that you will find the person in a social situation
- Conditional probabilities give the probability of one event given that another event has occurred



## Conditional Probability

- The conditional probability of event B occurring given that event A has occurred is given by:
- p(B|A) = p(A and B) / p(A)
- The values of p(A and B) and p(A) can be easily gotten from the table of joint and marginal probabilities



## Conditional Probability

- What is the probability that a person likes cats given he or she likes statisitcs?
- p(likes cats | likes stats) = p(likes cats and likes stats) / p(likes stats)
- .167 / .222 = .75

| | Likes Cats (B) | Does Not Like Cats (not B) | Marginal Probability |
|---|---|---|---|
| Likes Stat (A) | 3<br>.167 | 1<br>.056 | 4<br>.222 |
| Does Not Like Stats (not A) | 10<br>.556 | 4<br>.222 | 14<br>.778 |
| Marginal Probability | 13<br>.722 | 5<br>.278 | 18 |

# Conditional Probability

- What is the probability that a person likes cats given he or she does not like statisitcs?
- p(likes cats | dislikes stats) = p(likes cats and dislikes stats) / p(dislikes stats)
- .556 / .778 = .714

| | Likes Cats (B) | Does Not Like Cats (not B) | Marginal Probability |
|---|---|---|---|
| Likes Stat (A) | 3<br>.167 | 1<br>.056 | 4<br>.222 |
| Does Not Like Stats (not A) | 10<br>.556 | 4<br>.222 | 14<br>.778 |
| Marginal Probability | 13<br>.722 | 5<br>.278 | 18 |

21

# The Multiplication Rule Revisited

- The multiplication rule given before applied only to two (or more) events that were independent of each other
- When the event are not independent, the multiplication rule must be revised to:
- p(A and B) = p(A) X p(B | A)

22

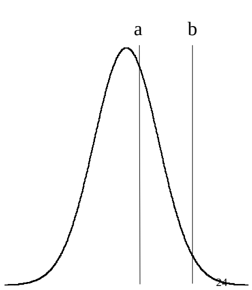
# Continuous Variables and Probability

- When the variables / events are continuous rather than discrete, we can no longer simply count the occurrence of the event or count the total number of events
- The continuous nature of the variable implies that there is an infinite number of values that the variable can take on

23

# Probability of Continuous Events

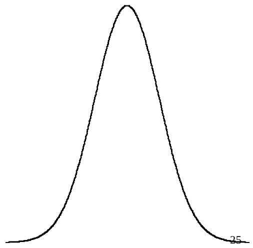
- Rather than counting, the probability is determined by areas under a probability curve
- p(a ≤ X ≤ b) = area under curve between a and b / total area under curve




24

# Unit Normal Curve

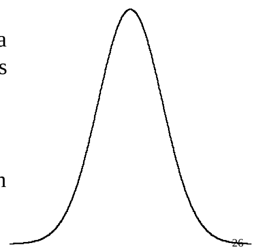
- The *unit normal curve* is frequently used in statistics to determine probabilities
- It is a normal (or Gaussian) curve with a standard deviation of 1
- The area under the unit normal curve is 1
- Many variables have approximately normal distributions





# Determining Probabilities with the Unit Normal Distribution

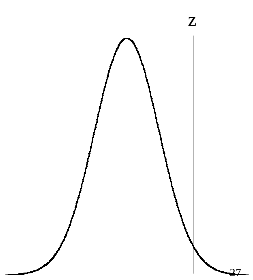
- Weights are normally distributed
- What is the probability of a randomly selected female's weight being larger than 190 pounds?
- The mean weight of females is 133 pounds with a standard deviation of 22 pounds





# Determining Probabilities with the Unit Normal Distribution

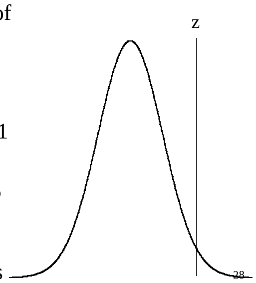
- Convert the raw score to a z-score:
- z = (190 - 133) / 22 = 2.59
- Use Table A in Appendix D of your text (or any other similar table) to find the area above a z-score of 2.59






# Determining Probabilities with the Unit Normal Distribution

- The area above a z-score of 2.59 is .0048
- The total area under the unit normal curve is 1
- The probability is .0048 / 1 = .0048
- There are approximately 5 chances in 1000 of a randomly selected female weighing over 190 pounds



## Probability Problems

- What is the probability of a randomly selected female weighing less than 100 pounds?

Solution

- What is the probability of a randomly selected female weighing more than 150 pounds or less than 110 pounds?

Solution

- What is the probability of two randomly selected females both weighing less than 90 pounds?

Solution

29

*The End*

30

## Draw a King

- How many kings are in a standard deck of cards?
  - 4
- How many cards in a deck of cards?
  - 52
- The probability of drawing a king from a deck is given by # kings / # cards = 4 / 52 = .076923

Problems

31

## Draw a Face Card

- There are three types of face cards: Jacks, Queens, and Kings
- Each face card is mutually exclusive
  - If a card is a Jack, it cannot be a Queen or King
- To determine the probability of one or more mutually exclusive events occurring, use the addition rule

32

## Draw a Face Card

- p(Jack or Queen or King) = p(Jack) or p(Queen) or p(King)
- From the previous problem, p(King) = .076923
- p(Jack) = p(Queen) = p(King)
- p(Jack or Queen or King) = .076923 + .076923 + .076923 = .230769



## Draw a Face Card

- Another way of looking at this problem is simply to count the number of face cards in a deck
  - There are 12 face cards in a deck
- p(face card) = # face cards / # cards in deck
  = 12 / 52 = .230769
- The two answers agree

Problems



## Draw Four Aces

- The card that is dealt first has no influence on which card is dealt second, third, or fourth (because the card is replaced before the next draw)
- The card that is dealt second has no influence on which card is dealt first, third, or fourth, and so on
- Thus the draws are independent events



## Draw Four Aces

- To determine the probability of several events happening together, use the multiplication rule
- p(Ace on first draw and Ace on second draw and Ace on third draw and Ace on fourth draw) = p(Ace on first draw) X p(Ace on second draw) X p(Ace on third draw) X p(Ace on fourth draw)

## Draw Four Aces

- p(Ace on first draw) = # of aces / # cards
- There are four aces and 52 cards in the deck
- p(Ace on first draw) = 4 / 52 = .076923
- The card is replaced and the deck is reshuffled
- p(Ace on second draw) = p(Ace on third draw) = p(Ace on fourth draw) = 4 / 52 = .076923



## Draw Four Aces

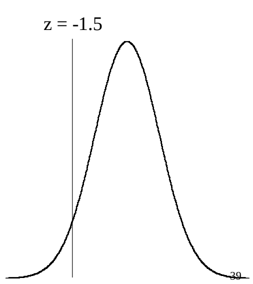
- p(draw four aces) = .076923 X .076923 X .076923 X .076923 = .00003501278
- This event would occur roughly 35 in 1,000,000 times

Problems



## Less Than 100 Pounds

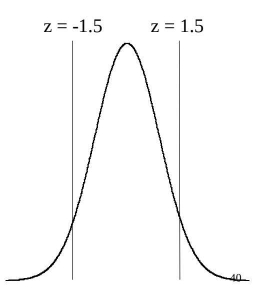
- First, convert the raw score to a z-score
- N(133, 22), raw score = 100
- z = (100 - 133) / 22 = -1.5
- Draw a unit normal distribution and put the z-score on it





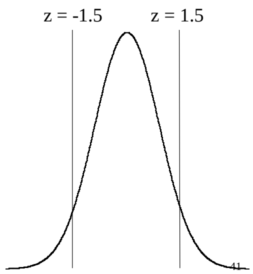

## Less Than 100 Pounds

- We want the area to the left of the z-score
- Because the unit normal distribution is symmetrical, the area to the left of z = -1.5 is the same as the area to the right of z = 1.5



## Less Than 100 Pounds

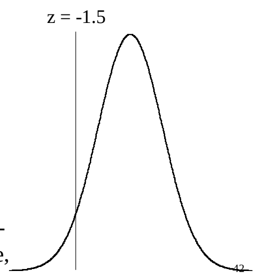
- Find the area to the right of z = 1.5
- Consult Table A or this table
- The area corresponds to .0668
- Thus, 6.68% of randomly selected females should weigh less than 100 pounds






## Less Than 100 Pounds

- Because the total area under the curve is 1, the area to the left of any z-score plus the area to the right of the z-score must equal 1
- Because of symmetry, the area to the left of a negative z-score equals 1 - area to the left of the same, positive z-score

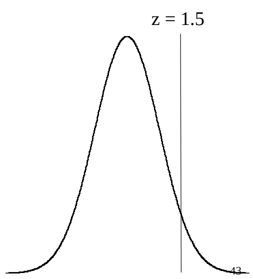





## Less Than 100 Pounds

- Area to the left of z equal to -1.5 equals 1 - area to the left of z equal 1.5
- 1 - .9332 = .0668

Problems



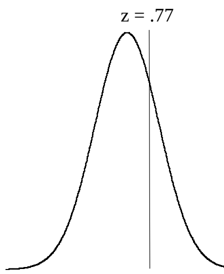



## > 150 or < 110

- This problem calls for the addition rule of mutually exclusive events
- Determine the probability of a randomly selected female weighing more than 150 pounds
- Determine the probability of a randomly selected female weighing less than 110 pounds

## > 150 Pounds

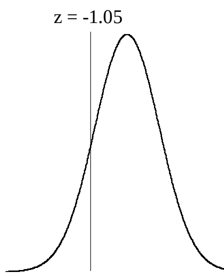
- Convert 150 pounds to a z-score
- z = 150 - 133 / 22 = 0.77
- Draw the distribution and z-score
- Consult a table to find the area above a z-score of .77
- p(> 150) = .2206






## < 110 Pounds

- Convert 110 pounds to a z-score
- z = 110 - 133 / 22 = -1.05
- Draw the distribution and z-score
- Consult a table to find the area below a z-score of -1.05
- p(< 110) = .1469






## p(>150 or <110)

- Combine the probabilities with the addition rule:
- p(>150 or <110) = p(>150) + p(<110) = .2206 + .1469 = .3675
- Roughly 37% of randomly selected women will weigh more than 150 pounds or less than 110 pounds

Problems

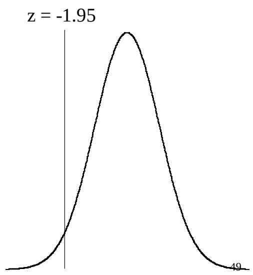


## Two Weighing Less than 90 Pounds

- This is an example of the multiplication rule for independent events
- p(A < 90 and B < 90) = p(A < 90) X p(B < 90)
- Determine the probability of randomly selecting one female weighing less than 90 pounds

## < 90

- Convert to a z-score:
- $z = (90 - 133) / 22 = -1.95$
- Draw the distribution and z-score
- Find, in a table, the area above z = 1.95
- $p(A < 90) = .0256$
- $p(B < 90) = .0256$






## Combine Using Multiplication Rule

- $p(A < 90 \text{ and } B < 90) = p(A < 90) \times p(B < 90)$
- $= .0256 \times .0256 = .0006554$
- Only about 66 times out of 100,000 times that you randomly selected two women would you expect both to weigh less than 90 pounds.

Problems