Lecture 11:
Lesson and Activity Packet

MATH 232: Introduction to Statistics

October 26, 2016

Homework and Announcements

- Homework 12 due in class on Friday
- Election Prediction Project
  - Friday will be a project day
  - Deadline for ASA submission is Sunday at 8 p.m.
  - Deadline for group report, proof of ASA submission, and individual report is Monday at 11:59 p.m.
  - Please let me know before Sunday whether your group wants to present in class on Monday
- Book Review Project
  - Details to be given after the Election Project, **but** you should buy or procure (by legal means...) the book of your choice now, from among the following three choices:
    1. *Moneyball: The Art of Winning an Unfair Game*, by Michael Lewis
    2. *Dataclysm: Love, Sex, Race, and Identity—What Our Online Lives Tell Us about Our Offline Selves*, by Christan Rudder
    3. *How Not to Be Wrong*, by Jordan Ellenberg

Last time

- Conditional probability

Recap of today

- More conditional probability
- Independent events
- Multiplication rules

Recall the definition of conditional probability: $P(E|F)$ is the probability that E will occur (or will be discovered to have occurred), given that F has occurred (or has been discovered to have occurred). $P(E|F)$ can be computed according to the following rule:

Definition 1

If $P(F) > 0$, then

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}.$$

All of the examples we saw in Packet 10 were derived from classical probability. Let's try some examples with a relative frequency definition.

Lecture 11:
Lesson and Activity Packet

MATH 232: Introduction to Statistics

October 26, 2016

Homework and Announcements

- Homework 12 due in class on Friday
- Election Prediction Project
  - Friday will be a project day
  - Deadline for ASA submission is Sunday at 8 p.m.
  - Deadline for group report, proof of ASA submission, and individual report is Monday at 11:59 p.m.
  - Please let me know before Sunday whether your group wants to present in class on Monday
- Book Review Project
  - Details to be given after the Election Project, **but** you should buy or procure (by legal means...) the book of your choice now, from among the following three choices:
    1. *Moneyball: The Art of Winning an Unfair Game*, by Michael Lewis
    2. *Dataclysm: Love, Sex, Race, and Identity—What Our Online Lives Tell Us about Our Offline Selves*, by Christan Rudder
    3. *How Not to Be Wrong*, by Jordan Ellenberg

Last time

- Conditional probability

Recap of today

- More conditional probability
- Independent events
- Multiplication rules

Recall the definition of conditional probability: $P(E|F)$ is the probability that E will occur (or will be discovered to have occurred), given that F has occurred (or has been discovered to have occurred). $P(E|F)$ can be computed according to the following rule:

Definition 1

If $P(F) > 0$, then

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}.$$

All of the examples we saw in Packet 10 were derived from classical probability. Let's try some examples with a relative frequency definition.

Example 1

Suppose that a consumer research organization has studied the service under warranty provided by the 200 tire dealers in a large city, and that their findings are summarized in the following table:

	Good service under warranty	Poor service under warranty	Total
Name-brand tire dealers	64	16	80
Off-brand tire dealers	42	78	120
Total	106	94	200

If one of these tire dealers is randomly selected (that is, each one has the probability $\frac{1}{200}$ of being selected), the probability of choosing a name-brand dealer is $P(N) :=$________. The probability of choosing a dealer who provides good service under warranty is $P(G) :=$________. The probability of choosing a name-brand dealer who provides good service under warranty is $P(N \text{ and } G) :=$________.

$P(N) = \frac{80}{200} = \frac{40}{100} = 0.4 = 40\%$

$P(G) = \frac{106}{200} = \frac{53}{100} = 0.53 = 53\%$

$P(N \cap G) = \frac{64}{200} = \frac{32}{100} = 32\%$

Notice: $P(G^c) = 1 - 0.53 = 0.47$. A 47% chance of getting bad service when selecting at random.

But $P(G|N) = \frac{\#\text{good service name-brand}}{\#\text{name-brand}} = \frac{64}{80} = \frac{8}{10} = 80\%$.

or $P(G|N) = \frac{P(N \cap G)}{P(N)} = \frac{64/200}{80/200} = \frac{64}{200} \cdot \frac{200}{80} = \frac{64}{80} = 80\%$.

If picking just from name-brand, then 80% chance of getting good service.

Definition 2 (*Independent Events*)

If $P(B) \neq 0$ and if $P(A|B) = P(A)$, we say that event A is independent *of event B. That is, event A is independent of event B if the probability of A is not affected by the occurrence or non-occurrence of event B.*

Group Exercise 1

The probabilities that it will rain or snow in a given city on Christmas Day, on New Year's Day, or on both days, are $P(C) := 0.60$, $P(N) := 0.60$, and $P(C \cap N) := 0.42$. Check whether N and C are independent.

$$P(N|C) = \frac{P(C \cap N)}{P(C)} = \frac{0.42}{0.6} = 0.7 \neq 0.6 = P(N),$$

So, no, not independent events.

Actually, $P(N|C) = 0.7 > 0.6 = P(N)$, so the chance of it snowing on New Year's Day is greater if it's already snowed on Christmas.

Some interesting Multiplication rules follow from

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}.$$

A few are:

Theorem 1

General multiplication rule

$$P(E \cap F) = P(F) \cdot P(E|F)$$

Theorem 2

General multiplication rule

$$P(E \cap F) = P(E) \cdot P(F|E)$$

If A and B are independent, then:

Theorem 3

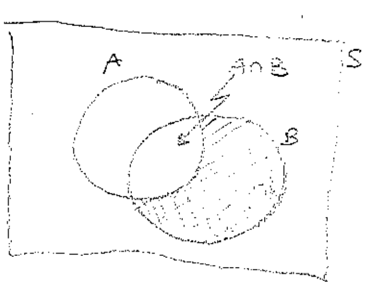
$$P(A \text{ and } B) = P(A) \cdot P(B)$$

See Packet 12 for examples.

Notes: If A & B independent, then $P(A) = \frac{P(A \cap B)}{P(B)}$

and $P(A^c) = 1 - P(A) = 1 - \frac{P(A \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)} = \frac{P(A^c \cap B)}{P(B)}$.

So A^c & B are independent too. Another useful identity.



The shaded region is $B - (A \cap B)$, which is the same as $A^c \cap B$.