Independence, Concentration, Bayes’ Theorem

CSE21 Winter 2017, Day 23 (B00), Day 15 (A00)

March 10, 2017

http://vlsicad.ucsd.edu/courses/cse21-w17
Independent Events

Two events $E$ and $F$ are **independent** iff $P(E \cap F) = P(E) P(F)$.

**Problem:** Suppose

- $E$ is the event that a randomly generated bitstring of length 4 starts with a 1
- $F$ is the event that this bitstring contains an even number of 1s.

Are $E$ and $F$ independent if all bitstrings of length 4 are equally likely? Are they disjoint?

**First impressions?**

A. $E$ and $F$ are independent and disjoint.
B. $E$ and $F$ are independent but not disjoint.
C. $E$ and $F$ are disjoint but not independent.
D. $E$ and $F$ are neither disjoint nor independent.
Independent Events

Two events E and F are independent iff \( P( E \cap F ) = P(E) P(F) \).

**Problem:** Suppose
- E is the event that a randomly generated bitstring of length 4 starts with a 1
- F is the event that this bitstring contains an even number of 1s.

Are E and F independent if all bitstrings of length 4 are equally likely? Are they disjoint?
Independent Random Variables

Let $X$ and $Y$ be random variables over the same sample space. $X$ and $Y$ are called independent random variables if, for all possible values of $v$ and $u$,

$$P(X = v \text{ and } Y = u) = P(X = v) P(Y = u)$$

Which of the following pairs of random variables on sample space of sequences of H/T when coin is flipped four times are independent?

A. $X_{12} = \# \text{ of H in first two flips}, \ X_{34} = \# \text{ of H in last two flips.}$
B. $X = \# \text{ of H in the sequence}, \ Y = \# \text{ of T in the sequence.}$
C. $X_{12} = \# \text{ of H in first two flips}, \ X = \# \text{ of H in the sequence.}$
D. None of the above.
Independence

Theorem:
If X and Y are independent random variables over the same sample space, then

\[ E(XY) = E(X)E(Y) \]

Note: This is not necessarily true if the random variables are not independent!
Concentration

How close (on average) will we be to the average / expected value?
Let \( X \) be a random variable with \( E(X) = E \).

The **unexpectedness** of \( X \) is the random variable
\[
U = |X - E|
\]
The **average unexpectedness** of \( X \) is
\[
AU(X) = E ( |X - E| ) = E(U)
\]
The **variance** of \( X \) is
\[
V(X) = E( |X - E|^2 ) = E(U^2)
\]
The **standard deviation** of \( X \) is
\[
\sigma(X) = ( E( |X - E|^2 ) )^{1/2} = V(X)^{1/2}
\]

*Rosen Section 7.4*
Concentration

How close (on average) will we be to the average / expected value?
Let $X$ be a random variable with $E(X) = E$.

The **unexpectedness** of $X$ is the random variable
$$U = |X - E|$$

The **average unexpectedness** of $X$ is
$$AU(X) = E(|X - E|) = E(U)$$

The **variance** of $X$ is
$$V(X) = E(|X - E|^2) = E(U^2)$$

The **standard deviation** of $X$ is
$$\sigma(X) = \sqrt{E(|X - E|^2)} = \sqrt{V(X)}$$

*Weight all differences from mean equally*

*Weight large differences from mean more*
Concentration

How close (on average) will we be to the average / expected value?
Let \( X \) be a random variable with \( E(X) = E \).

The unexpectedness of \( X \) is the random variable
\[
U = |X - E|
\]
The **average unexpectedness** of \( X \) is
\[
AU(X) = E(|X - E|) = E(U)
\]
The variance of \( X \) is
\[
V(X) = E(|X - E|^2) = E(U^2)
\]
The **standard deviation** of \( X \) is
\[
\sigma(X) = (E(|X - E|^2))^{1/2} = V(X)^{1/2}
\]

Does the average unexpectedness equal the standard deviation?

A. Yes
B. No
C. ??
Concentration

How close (on average) will we be to the average / expected value?
Let \( X \) be a random variable with \( E(X) = E \).
The variance of \( X \) is

\[
V(X) = E( |X - E|^2 ) = E(U^2)
\]

Example: \( X_1 \) is a random variable with distribution
\[
P( X_1 = -2 ) = 1/5, \ P( X_1 = -1 ) = 1/5, \ P( X_1 = 0 ) = 1/5, \ P( X_1 = 1 ) = 1/5, \ P( X_1 = 2 ) = 1/5.
\]

\( X_2 \) is a random variable with distribution
\[
P( X_2 = -2 ) = 1/2, \ P( X_2 = 2 ) = 1/2.
\]

Which is true?
A. \( E(X_1) \neq E(X_2) \)
B. \( V(X_1) < V(X_2) \)
C. \( V(X_1) > V(X_2) \)
D. \( V(X_1) = V(X_2) \)
Concentration

How close (on average) will we be to the average / expected value?

Let $X$ be a random variable with $E(X) = E$.

The variance of $X$ is

$$V(X) = E(|X - E|^2) = E(U^2)$$

*Theorem*: $V(X) = E(X^2) - (E(X))^2$
Concentration

How close (on average) will we be to the average / expected value?
Let $X$ be a random variable with $E(X) = E$.
The variance of $X$ is

$$V(X) = E(|X - E|^2) = E(U^2)$$

**Theorem:** $V(X) = E(X^2) - (E(X))^2$

**Proof:**

$$V(X) = E((X-E)^2) = E(X^2 - 2XE + E^2) = E(X^2) - 2E E(X) + E^2$$

$$= E(X^2) - 2E^2 + E^2$$

**Linearity of expectation**

$$= E(X^2) - (E(X))^2$$

😊
Concentration

How close (on average) will we be to the average / expected value?
Let \( X \) be a random variable with \( E(X) = E \).
The variance of \( X \) is

\[
V(X) = E(|X - E|^2) = E(U^2)
\]

**Theorem:** \( V(X) = E(X^2) - (E(X))^2 \)

What does it mean for the variance of \( X \) to be 0?

- A. The value of the random variable \( X \) must always be 0.
- B. The value of the random variable must be constant.
- C. \( E(X^2) = E(X) \)
- D. \( E(X^2) = E(U^2) \)
- E. \( E(U) = 0 \)
Recall: Conditional probabilities

Probability of an event may **change** if we have additional information about outcomes.

Suppose $E$ and $F$ are events, and $P(F) > 0$. Then,

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

i.e.,

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

*Rosen p. 456*
Bayes' Theorem

\[ P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})} \]

Based on previous knowledge about how probabilities of two events relate to one another, how does knowing that one event occurred impact the probability that the other event occurred?
A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Your first guess?
A. Close to 95%
B. Close to 85%
C. Close to 50%
D. Close to 15%
E. Close to 10%
A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Define events: we want $P(\text{used steroids} | \text{tested positive})$
A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Define events: we want \( P(\text{used steroids} | \text{tested positive}) \) so let

\[
E = \text{Tested positive} \\
F = \text{Used steroids}
\]
Bayes' Theorem: Example 1

Rosen Section 7.3

A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Define events: we want \( P( \text{used steroids} \mid \text{tested positive}) \)

E = Tested positive \( P( E \mid F ) = 0.95 \)
F = Used steroids
A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Define events: we want $P(\text{ used steroids } \mid \text{ tested positive})$

$$E = \text{Tested positive} \quad P(E \mid F) = 0.95$$

$$F = \text{Used steroids} \quad P(F) = 0.1 \quad P(F^c) = 0.9$$
Bayes' Theorem: Example 1

A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Define events: we want \( P( \text{used steroids} \mid \text{tested positive}) \)

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid \overline{F})P(\overline{F})}
\]

\[
E = \text{Tested positive} \quad P( E \mid F ) = 0.95 \quad P( E \mid \overline{F} ) = 0.15
\]

\[
F = \text{Used steroids} \quad P(F) = 0.1 \quad P(\overline{F}) = 0.9
\]
Bayes' Theorem: Example 1

A manufacturer claims that its drug test will detect steroid use 95% of the time. What the company does not tell you is that 15% of all steroid-free individuals also test positive (the false positive rate). 10% of the Tour de France bike racers use steroids. Your favorite cyclist just tested positive. What’s the probability that he used steroids?

Define events: we want \( P( \text{used steroids} \mid \text{tested positive}) \)

\[
P(E \mid F) = 0.95 \quad P(E \mid F') = 0.15
\]

\[
P(F) = 0.1 \quad P(F') = 0.9
\]

Plug in:

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F')P(F')} = \frac{0.95 \cdot 0.1}{0.95 \cdot 0.1 + 0.15 \cdot 0.9} = 0.41
\]

Plug in: 41%
Bayes' Theorem: Example 2

Suppose we have found that the word “Rolex” occurs in 250 of 2000 messages known to be spam and in 5 out of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam.

\[
P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}
\]

Your first guess?

A. Close to 95%
B. Close to 85%
C. Close to 50%
D. Close to 15%
E. Close to 10%
Bayes' Theorem: Example 2

Rosen Section 7.3

\[
P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}
\]

Suppose we have found that the word “Rolex” occurs in 250 of 2000 messages known to be spam and in 5 out of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam.

We want: \( P(\text{spam} \mid \text{contains "Rolex"}) \). So define the events

\( E = \text{contains "Rolex"} \)
\( F = \text{spam} \)
Bayes' Theorem: Example 2

Suppose we have found that the word “Rolex” occurs in 250 of 2000 messages known to be spam and in 5 out of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam.

We want: \( P(\text{spam} \mid \text{contains "Rolex"}) \). So define the events

\[
E = \text{contains "Rolex"}
\]

\[
F = \text{spam}
\]

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid \bar{F})P(\bar{F})}
\]

What is \( P(E \mid F) \)?

A. 0.005  
B. 0.125  
C. 0.5  
D. Not enough info
Suppose we have found that the word “Rolex” occurs in 250 of 2000 messages known to be spam and in 5 out of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam.

We want: \( P(\text{spam} \mid \text{contains } \text{"Rolex"}) \).

\[
P(F\mid E) = \frac{P(E\mid F)P(F)}{P(E\mid F)P(F) + P(E\mid \overline{F})P(\overline{F})}
\]

\( E = \text{contains } \text{"Rolex"} \) \hspace{1cm} \( P(E\mid F) = \frac{250}{2000} = 0.125 \) \hspace{1cm} \( P(E\mid \overline{F}) = \frac{5}{1000} = 0.005 \)
Bayes' Theorem: Example 2

\[ P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})} \]

Suppose we have found that the word “Rolex” occurs in 250 of 2000 messages known to be spam and in 5 out of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam.

We want: \( P(\text{spam} | \text{contains "Rolex"}) \).

\[ E = \text{contains “Rolex”} \quad P( E | F ) = 250/2000 = 0.125 \quad P( E | F' ) = 5/1000 = 0.005 \]
\[ F = \text{spam} \quad P( F ) = P( F' ) = 0.5 \]
Bayes' Theorem: Example 2

Suppose we have found that the word “Rolex” occurs in 250 of 2000 messages known to be spam and in 5 out of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam.

We want: \( P( \text{spam} \mid \text{contains } "\text{Rolex}" ) \).

\[
P(\overline{F} \mid E) = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid \overline{F})P(\overline{F})} = \frac{0.125 \cdot 0.5}{0.125 \cdot 0.5 + 0.005 \cdot 0.5} = 0.96
\]

Plug in: 96%
Announcements

OHs and 1-1 sessions
Plenty available!

HW #8
due Tuesday, March 14
11:59pm