Topic 1: Probability

STAT 211 - 509

9/4/2018

Sample spaces and events

- **Outcome**: possible result of an experiment.
- **Sample space** (S): set of all possible outcomes of an experiment.
- **Event**: any collection (subset) of outcomes contained in the sample space S.

Example

Toss a fair coin twice. List the sample space and define the event A as "the first toss is heads"

- Sample space: $S = \{HH, HT, TH, TT\}$
- Event: $A = \{HH, HT\}$

Set operations

- The **union** of two events A and B, denoted $A \cup B$ is the event consisting of all outcomes that are either in A, in B, or in both
- The **intersection** of two events A and B, denoted $A \cap B$ is the consisting of all outcomes in both A and B. A and B are **mutually exclusive** if they do not have any outcomes in common.
- The **complement** of an event A, denoted A', is the set of all outcomes in S that are not contained in A

Probability

- Probability theory is the study of randomness and uncertainty
- Given an experiment and a sample space S, want to assign each A the number $P(A)$, which is the measure of the chance that A will occur
- The assignments $P(\cdot)$ should satisfy axioms that accord with intuitive notions of probability

Axioms of probability

1. For any event A, $P(A) \geq 0$
2. $P(S) = 1$
3. If $A_1, A_2, \ldots, A_k$ is a collection of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_k) = \sum_{i=1}^{k} P(A_i)$$

Implications

- For the empty set $\emptyset$, $P(\emptyset) = 0$
- For any event A:
  - $0 \leq P(A) \leq 1$
  - $P(A') = 1 - P(A)$
- For any events A and B:

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

Example

Toss a fair coin twice. What is the probability that at least one toss is heads? First, define relevant events

- Sample space: $S = \{HH, HT, TH, TT\}$
- Let $A = \{\text{at least one toss is heads}\} = HH \cup HT \cup TH$

  What is $P(A)$?
- Union:

$$P(A) = P(HH \cup HT \cup TH) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

- Complement:

$$\begin{aligned} A' &= \{\text{no toss is heads}\} = TT \\ P(A') &= \frac{1}{4} \\ P(A) &= 1 - P(A') = \frac{3}{4} \end{aligned}$$

- Another:

Let $H_1 = \{\text{heads on first toss}\}$ and $H_2 = \{\text{heads on second toss}\}$
Then $A = \{\text{at least one toss is heads}\} = H_1 \cup H_2$

$$\begin{aligned} P(A) &= P(H_1 \cup H_2) \\ &= P(H_1) + P(H_2) - P(H_1 \cap H_2) \\ &= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} \end{aligned}$$

What is the probability exactly one toss is heads?
$H_1 = \{\text{heads on first toss}\}$ and $H_2 = \{\text{heads on second toss}\}$
$A = \{\text{exactly one toss is heads}\} = (H_1 \cap H_2') \cup (H_1' \cap H_2)$

Uniform probability

Suppose we have a finite sample space S with N outcomes, integer $N \geq 1$. Each outcome is equally likely. If there are m outcomes in event A, then

$$P(A) = \frac{m}{N}$$

Example

Roll two fair dice, each outcome in sample space has equal probability $1/36$

$$\begin{aligned} S = \{ & \\ & (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ & (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ & (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ & (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ & (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ & (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \\ & \} \end{aligned}$$

What is the probability that the sum of the two faces is 7?
$A = \{\text{sum of the two faces is 7}\} = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$
6 outcomes in A, 36 possible outcomes: $P(A) = \frac{6}{36} = \frac{1}{6}$

Counting

Suppose we toss a coin 10 times.

- How many outcomes are there?
- How many would have exactly 4 heads?
- What is the probability of exactly 4 heads?

Counting rules

- Rule of sum: if there are a ways to do thing one and b ways to do thing two, and you can't do both things, then there are $a + b$ things to do in total
- Rule of product: if there are a ways to do thing one, b ways to do thing two, then there are ab ways to do both

Binomial coefficient

- **Binomial coefficients** are family of positive integers that occur as coefficients in the binomial theorem
- Coefficient of the x^k term in the polynomial expansion of $(1 + x)^n$
- Indexed by two nonnegative integers n and k
- $$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$
- The number of distinct ways of choosing k objects out of n

Example

We toss a coin 10 times

- How many possible outcomes are there?

  2^{10}

- How many of these outcomes have exactly 4 heads?

  $\binom{10}{4}$

```r
choose(n = 10, k = 4)
```

```
## [1] 210
```

- What is the probability of exactly 4 heads?

  $P(X = 4) = \frac{\binom{10}{4}}{2^{10}} = \frac{210}{1024} = 0.205078125$

Independence

- Two events A and B are **independent** if and only if

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

- Intuitively, probability of A does not depend on whether or not B occurred, and vice versa.

- Many statistical methods assume independence between all pairs of observations in dataset

- Independence of two events does not imply they are mutually exclusive

- Examples of dependence:

  - Multiple measurements of same individual over time
  - Measurements located near each other

NFL example

Are the outcomes of two games independent of one another?

- Suppose previous game was demoralizing loss. Is the team extra motivated for next?
- Suppose the team won previous game, guaranteeing entrance into playoffs. Is the team *less* motivated for next?
- Suppose team has won 5 games in a row. Do they have momentum?

Random variables and distributions

- For a sample space S of some experiment, a **random variable (rv)** is any rule that associates a number with each outcome in S.

- Etymology:

  - called "variable" because different numerical values possible
  - "random" because observed value depends on uncertain experimental outcome

- Function that maps from sample space to real numbers: $X : S \to \mathbb{R}$

- Random variables have probability distributions that specify the probability of the rv falling in an interval

Why random variables

- Leads to easier math and calculations

- Can define multiple random variables on same probability space

- In statistics we usually care about distributions of random variables, not sample space

Types of random variables

- A random variable is **discrete** if its possible values are from a finite set, or can be listed in an infinite sequence with first, second, etc. elements
- A random variable is **continuous** if its possible values are from an entire interval of the real line
- The random variables $X_1, X_2, \ldots, X_n$ are **independent and identically distributed (iid)** if they are mutually independent and all follow the same distribution

Probability for discrete random variables

- The **probability mass function (pmf)** of a discrete rv is defined for every number x by

$$f(x) = P(X = x) = P(s \in S : X(s) = x)$$

$P(X = x)$ is the probability that the rv X assumes the value x

- The **cumulative distribution function (cdf)** $F(x)$ of a discrete rv X with pmf $f(x)$ is defined for every number x by

$$F(x) = P(X \leq x) = \sum_{y: y \leq x} f(y)$$

For any number x, $F(x)$ is the probability that the observed value of X will be at most x

Example

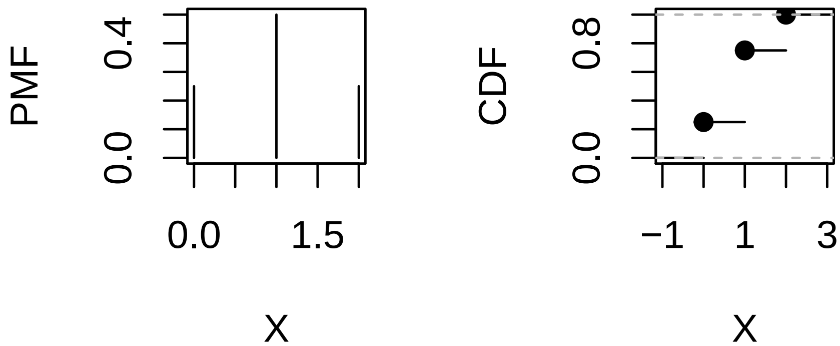
Flip a fair coin twice. Derive and plot the pmf and cdf. Let X be the random variable that equals the number of occurences of heads. The possible values of X are $0, 1, 2$.

- pmf:

$$P(X = x) = \begin{cases} 1/4, & x = 0 \\ 1/2, & x = 1 \\ 1/4, & x = 2 \\ 0, & x \notin \{0, 1, 2\} \end{cases}$$

- cdf:

$$P(X \leq x) = \begin{cases} 0, & x < 0 \\ 1/4, & 0 \leq x < 1 \\ 3/4, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$



The binomial distribution

Bernoulli distribution

- Consider experiment with single success/failure trial, with probability of success $p \in [0,1]$. The Bernoulli random variable X associated with this experiment is 0 if trial is a failure, 1 if it is a success.
- pmf of X:

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & x \in \{0,1\} \\ 0 & \text{otherwise} \end{cases}$$

- p is **parameter** of the distribution, must know to evaluate $f(x)$

Binomial distribution

- Consider an experiment consisting of n independent Bernoulli trails. The binomial random variable X associated with this experiment is defined as the number of successes out of the n trials
- pmf of X:

$$f(x) = \begin{cases} \binom{n}{x} p^x(1-p)^{1-x} & x \in 0,1,\ldots,n \\ 0 & \text{otherwise} \end{cases}$$

- n and p are parameters of the distribution
- We write $X \sim Binomial(n,p)$

Example

Suppose we toss a coin 10 times. $X \sim Binomial(10, 0.5)$. What is the probability of exactly 4 heads?

$$P(X=4) = f(4) = \binom{10}{4} 0.5^4 0.5^{10-4} = 0.205078125$$

```
dbinom(x = 4, size = 10, prob = 0.5)
```

```
## [1] 0.2050781
```

NFL example

- Let X be the number of wins out of a season's 16 games. If we assume that the outcome of each game is independent of all others, we can say $X \sim Binomial(16, p)$, where p is the probability that the Texans win any given game.
- Suppose that $p = 0.5$. What is the probability the Texans win 9 or more games in a season?
- $P(X \geq 9) = \sum_{x=9}^{16} f(x) = 0.4018097$

```
sum(dbinom(9:16, 16, 0.5))
1 - pbinom(q = 8, size = 16, prob = 0.5)
pbinom(q = 8, size = 16, prob = 0.5, lower.tail = FALSE)
```

- Would not be unusual at all for an "average" team to win 9 or more games in a season, just by chance.

Other discrete distributions

- **Discrete uniform distribution**: rv X with integer parameters a, b, $a \leq b$, with pmf:

$$f(x) = \frac{1}{b-a+1} \quad x = a, a+1, \ldots, b-1, b$$

- **Geometric distribution**: rv X with parameter $p > 0$, with pmf:

$$f(x) = (1-p)^x p \quad x = 0, 1, 2, \ldots$$

Describes the number of failures required until the first success in a series of Bernoulli trials

- **Poisson distribution**: rv with parameter $\lambda > 0$, with pmf:

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \ldots$$

Describes counts of events, like number of 911 calls on Friday night