03 - Probability

HCI/PSYCH 522
Iowa State University

January 25, 2022

Probability

What can you tell me about probability?

Kolmogorov's Axioms of Probability

1. Let E be an event, i.e. something happens. Then $P(E) \geq 0$.
2. Let Ω be the sample space, i.e. the collection of all possible outcomes. Then $P(\Omega) = 1$.
3. Let $E_1, E_2, \ldots$ be mutually exclusive events, i.e. no pair can both occur.

$$P(E_1 \text{ or } E_2 \text{ or } E_3 \ldots) = P(E_1) + P(E_2) + P(E_3) + \cdots .$$

Results from Kolmogorov's Axioms

Let A and B both be events.

- $0 \leq P(A) \leq 1$
- Let $\emptyset$ be the empty set, i.e. the "event" that nothing happens. $P(\emptyset) = 0$.
- Let A^C be the complement of A, i.e. all the outcomes that are not in A. Then $P(A^C) = 1 - P(A)$.
- Let $A \subset B$ indicate that A is a subset of B, i.e. all outcomes in B are also in A. Then $P(A) \leq P(B)$.
- $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

Conditional Probability

Let $P(A|B)$ indicate the conditional probability of A given B, i.e. we know B has occurred. Define

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \qquad P(B) > 0.$$

Independence

Two events A and B are independent if $P(A|B) = P(A)$ or, equivalently, $P(B|A) = P(B)$. Using the definition of conditional probability, we can find that for two independent events

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

Bayes' Rule

Bayes' Rule states

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^C)P(A^C)}$$

Diganostic Testing

If a pregnant woman has a test for Down syndrome and it is positive, what is the probability that the child will have Down syndrome? Let D indicate a child with Down syndrome and D^c the opposite. Let '$+$' indicate a positive test result and '$-$' a negative test result.

$$\begin{aligned}
\text{sensitivity} &= P(+|D) &= 0.94 \\
\text{specificity} &= P(-|D^c) &= 0.77 \\
\text{prevalence} &= P(D) &= 1/1000
\end{aligned}$$

$$\begin{aligned}
P(D|+) &= \frac{P(+|D)P(D)}{P(+)} = \frac{P(+|D)P(D)}{P(+|D)P(D)+P(+|D^c)P(D^c)} = \frac{0.94\cdot 0.001}{0.94\cdot 0.001+0.23\cdot 0.999} \\
&\approx 1/250
\end{aligned}$$

$$P(D|-) \approx 1/10,000$$

How do we interpret probability in the real world?

Relative frequency interpretation: probability is the proportion of times an event occurs in an infinite number of trials.

Personal belief interpretation: probability is a statement of how sure you are that an event will occur.

Random variables

Let ω be the outcome of an "experiment" (any data collection). Then $X(\omega) \in \mathbb{R}$ is a random variable, i.e. it is a function of the outcome of an "experiment" that returns a number.

We often know the following quantities for random variables:

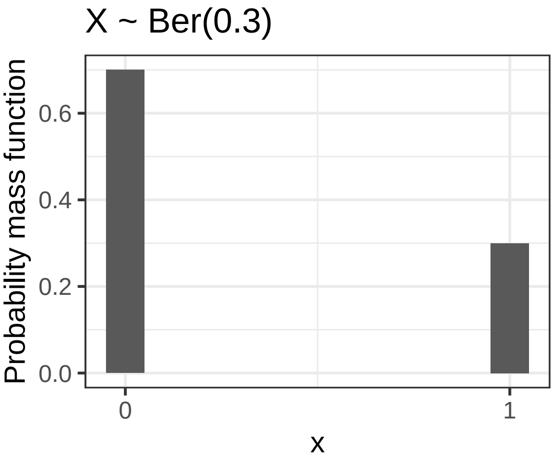
- Expectation (average value), $E[X]$
- Variance (variability), $Var[X]$
- Distribution:
  - Image, i.e. the possible values for X
  - Cumulative distribution function (cdf), $P(X \leq x)$
  - For discrete random variables, probability mass function (pmf) $P(X = x)$
  - For continuous random variables, probability density function (pdf) $f_X(x)$

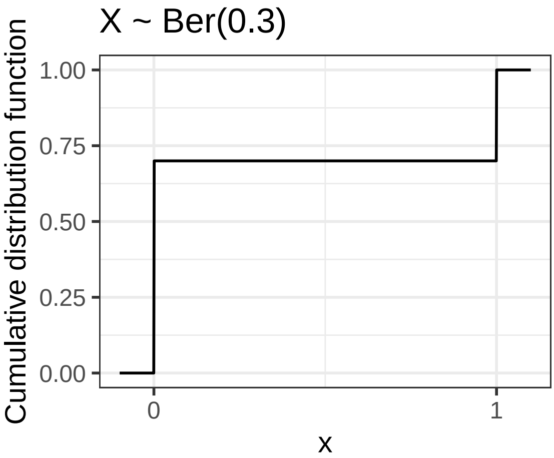
Bernoulli

If $X \sim Ber(p)$, then X is a Bernoulli random variable with probability of success p and

- $P(X = 1) = p$
- $P(X = 0) = (1 - p)$
- $E[X] = p$
- $Var[X] = p(1 - p)$

Bernoulli


X ~ Ber(0.3)
Probability mass function
0.6
0.4
0.2
0.0
0
1
x


X ~ Ber(0.3)
Cumulative distribution function
1.00
0.75
0.50
0.25
0.00
0
1
x

Binomial

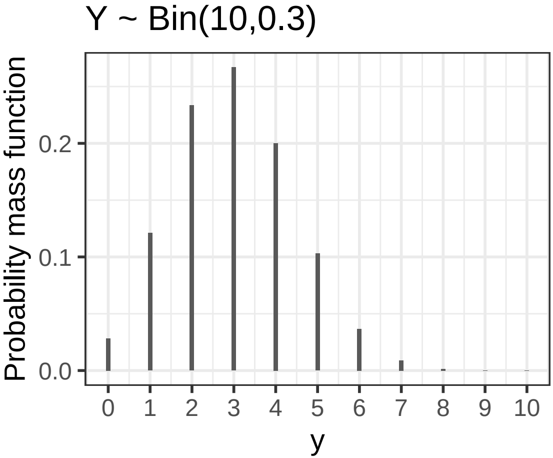
If $Y \sim Binom(n, p)$, then Y is a binomial random variable with n attempts and probability of success p and

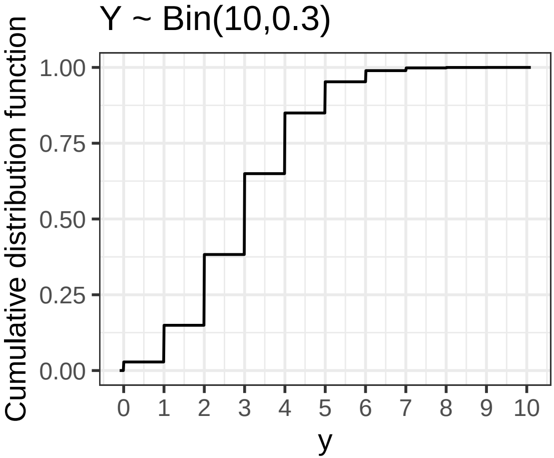
- $P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}$ for $y = 0, 1, \ldots, n$
- $E[Y] = np$
- $Var[Y] = np(1-p)$

```
n <- 10
p <- 0.3
dbinom(0:n, size = n, prob = p) %>% round(2)

##  [1] 0.03 0.12 0.23 0.27 0.20 0.10 0.04 0.01 0.00 0.00 0.00
```

Binomial


Y ~ Bin(10,0.3)
Probability mass function
0.0
0.1
0.2
0 1 2 3 4 5 6 7 8 9 10
y


Y ~ Bin(10,0.3)
Cumulative distribution function
0.00
0.25
0.50
0.75
1.00
0 1 2 3 4 5 6 7 8 9 10
y

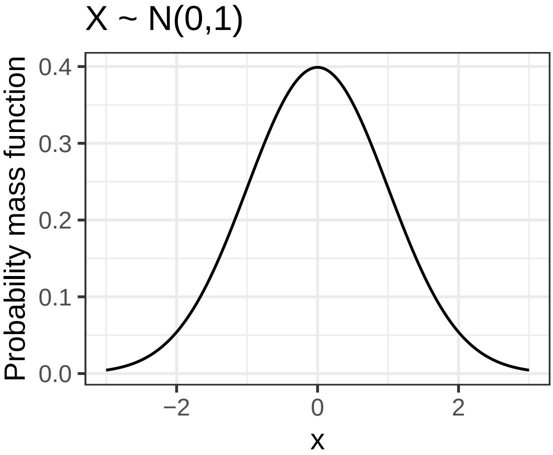
Normal

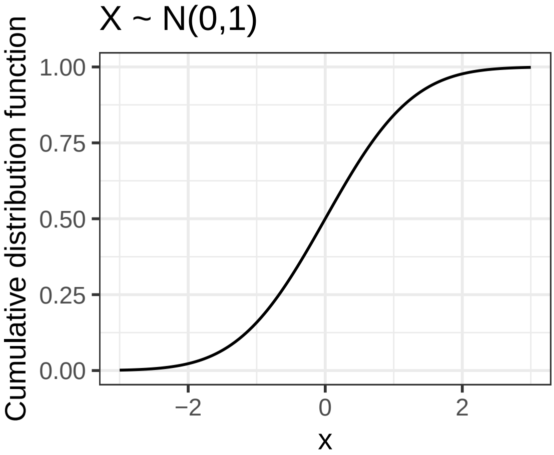
If $X \sim N(\mu, \sigma^2)$, then X is a normal random variable with mean μ and variance σ^2 (standard deviation σ) and

- $E[X] = \mu$
- $Var[X] = \sigma^2$ $(SD[X] = \sigma)$
- probability density function

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right).$$

Normal


X ~ N(0,1)
Probability mass function
0.4
0.3
0.2
0.1
0.0
−2
0
2
x


X ~ N(0,1)
Cumulative distribution function
1.00
0.75
0.50
0.25
0.00
−2
0
2
x