Introduction to Probability

Properties and distributions

Outline

- Properties of probability
  - Chain rule
  - Bayes Theorem
  - Expected values
- Common distributions
  - Discrete
  - Continuous

Chain Rule of Probability

- For two variables:

$$\Pr(X,Y) = \Pr(X \mid Y)P(Y)$$

- For three variables

$$\Pr(X,Y,Z) = \Pr(X \mid Y,Z)P(Y \mid Z)P(Z)$$

- etc.
- to see that this is true, note that

$$\Pr(X,Y,Z) = \frac{\Pr(X,Y,Z)}{P(Y,Z)}\frac{P(Y,Z)}{P(Z)}P(Z)$$

Bayes Theorem

$$\Pr(x \mid y) = \frac{\Pr(y \mid x)\Pr(x)}{\Pr(y)} = \frac{\Pr(y \mid x)\Pr(x)}{\sum_x \Pr(y \mid x)\Pr(x)}$$

- this theorem is extremely useful
- there are many cases when it is hard to estimate $\Pr(x \mid \mathrm{y})$ directly, but it's not too hard to estimate $\Pr(y \mid x)$ and $\Pr(x)$

Bayes Theorem Example

- MDs usually aren't good at estimating $\Pr(\mathit{Disorder} \mid \mathit{Symptom})$
- they're usually better at estimating $\Pr(\mathit{Symptom} \mid \mathit{Disorder})$
- if we can estimate $\Pr(\mathit{Fever} \mid \mathit{Flu})$ and $\Pr(\mathit{Flu})$ we can use Bayes' Theorem to do diagnosis

$$\Pr(\mathit{flu} \mid \mathit{fever}) = \frac{\Pr(\mathit{fever} \mid \mathit{flu}) \Pr(\mathit{flu})}{\Pr(\mathit{fever} \mid \mathit{flu}) \Pr(\mathit{flu}) + \Pr(\mathit{fever} \mid \neg \mathit{flu}) \Pr(\neg \mathit{flu})}$$

Expected Values

- the *expected value* of a random variable that takes on numerical values is defined as:

$$E[X] = \sum_x x \times \Pr(x)$$

this is the same thing as the *mean*

- we can also talk about the expected value of a function of a random variable (which is also a random variable)

$$E[g(X)] = \sum_x g(x) \times \Pr(x)$$

Expected Value Examples

$$E[Shoesize] =$$
$$5 \times \Pr(\mathrm{S}hoesize = 5) + \ldots + 14 \times \Pr(Shoesize = 14)$$

- Suppose each lottery ticket costs \$1 and the winning ticket pays out \$100. The probability that a particular ticket is the winning ticket is 0.001.

$$E[gain(Lottery)] =$$
$$gain(\text{winning}) \Pr(\text{winning}) + gain(\text{losing}) \Pr(\text{losing}) =$$
$$(\$100 - \$1) \times 0.001 - \$1 \times 0.999 =$$
$$-\$0.90$$

Linearity of Expectation

- An extremely useful aspect of expected values is the following identity

$$E[X+Y] = E[X] + E[Y]$$

- This holds even if X and Y are *not independent!*

Expected values of indicator random variables

- It is common to use *indicator* random variables

$$I_A = \begin{cases} 1 & \text{if event } A \text{ occurs,} \\ 0 & \text{otherwise} \end{cases}$$

- The expected value of such a variable is simply

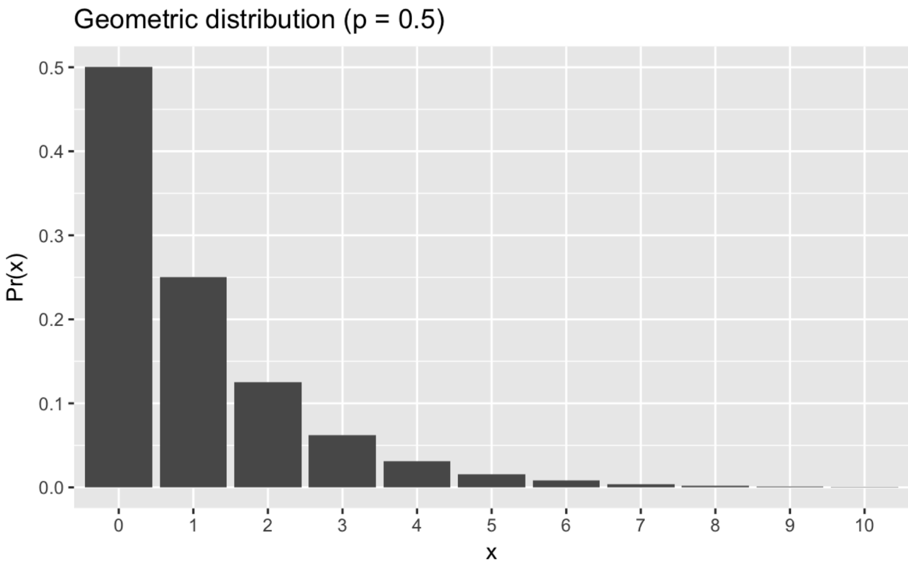
$$E[I_A] = 1 \times P(A) + 0 \times P(\neg A) = P(A)$$

The Geometric Distribution

- distribution over the number of trials before the first failure (with same probability of success p in each)

$$\Pr(x) = (1 - p)\, p^{x}$$

- e.g. the probability of x heads before the first tail

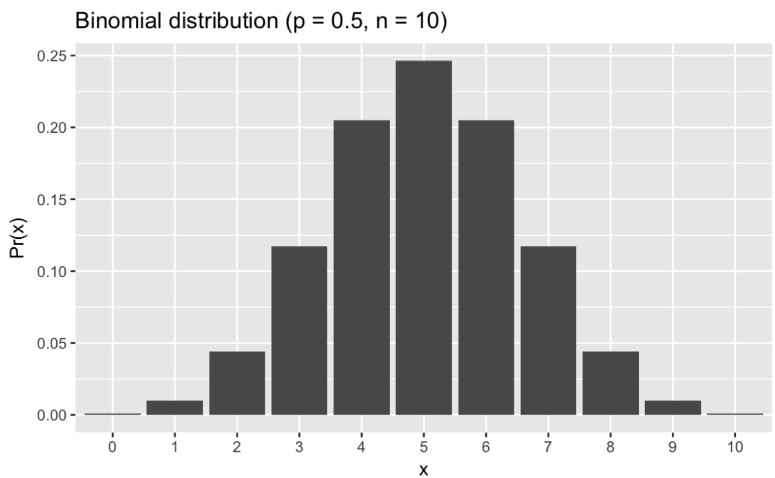


The Binomial Distribution

- distribution over the number of successes in a fixed number n of independent trials (with same probability of success p in each)

$$\Pr(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- e.g. the probability of x heads in n coin flips



The Multinomial Distribution

- k possible outcomes on each trial
- probability p_i for outcome i in each trial
- distribution over the number of occurrences x_i for each outcome in a fixed number n of independent trials

$$\Pr(x) = \frac{n!}{\prod_i (x_i!)} \prod_i p_i^{x_i}$$

For example, with $k = 6$ (e.g., a six-sided die) and $n = 30$:

$$\Pr([7,3,0,8,10,2]) = \frac{30!}{7!\times 3!\times 0!\times 8!\times 10!\times 2!}\left(p_1{}^7 p_2{}^3 p_3{}^0 p_4{}^8 p_5{}^{10} p_6{}^2\right)$$

Continuous random variables

- When our outcome is a continuous number we need a continuous random variable
- Examples: Weight, Height
- We specify a density function for random variable *X* as

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

- Probabilities are specified over an interval to derive probability values

$$P(a < X < b) = \int_a^b f(x)dx$$

- Probability of taking on a single value is 0.

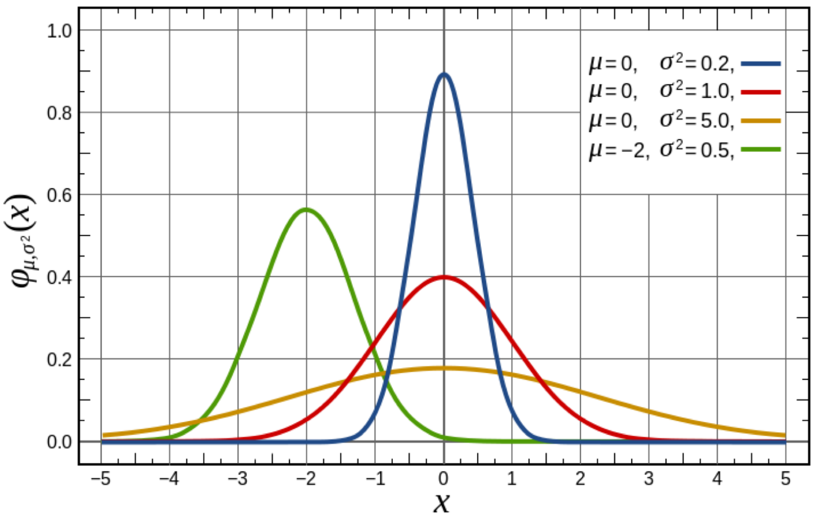
Examples of continuous distributions

- Gaussian distribution
- Exponential distribution
- Extreme Value distribution

Gaussian distribution

- Gaussian distribution

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



Summary

- Mathematics of distributions
  - Chain rule, Bayes Theorem
- Expected values
  - Linearity of expectation
- Distributions
  - Discrete: Geometric, Binomial, Multinomial
  - Continuous: Gaussian, Exponential, Extreme Value Distribution