Statistical Methods and Data Analysis I
Lecture 6: Probability Distributions

Oleg Goldshmidt
oleg.goldshmidt@post.idc.ac.il

Arison School of Business
Interdisciplinary Center (IDC)
Herzliya, Israel

March 19, 2019

Probability Distribution: Definition

Probability distribution

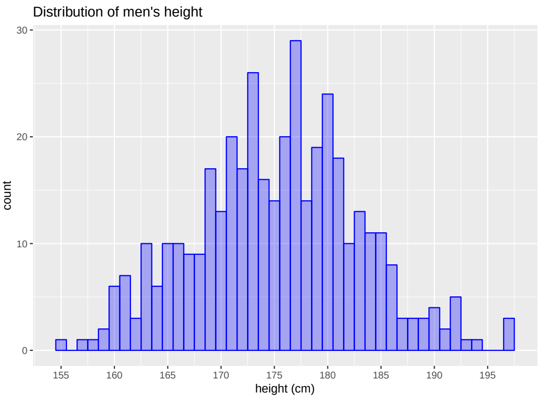
Lists all possible events and the probabilities with which they occur.

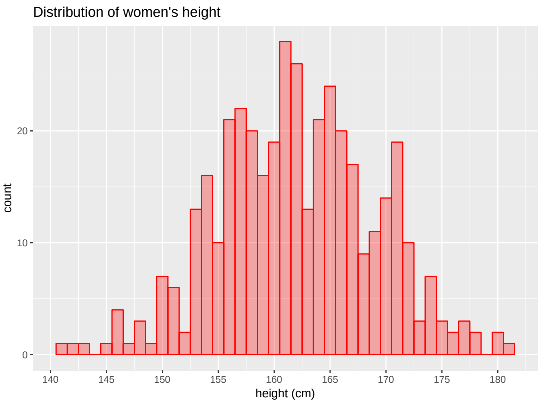
Rules:

1. The events must be disjoint.
2. All the probabilities must be between 0 and 1 inclusive.
3. The sum of all probabilities (the probability of *"something happening"*) must be 1.

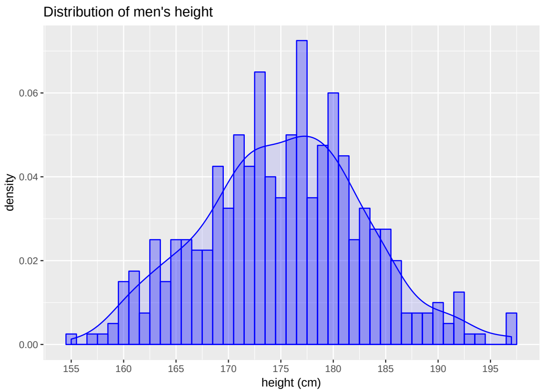
Easy enough if the events are discrete and their number is small enough.

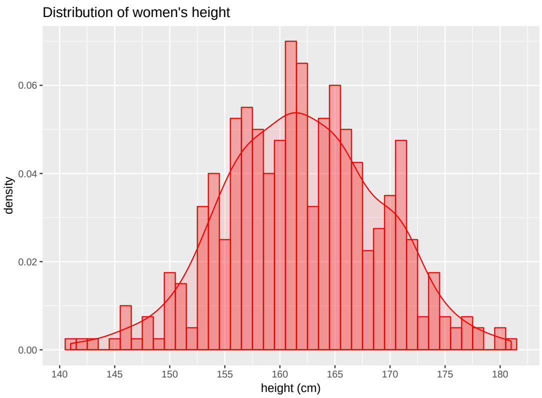
Visualizing Discrete Probability Distributions: Histograms



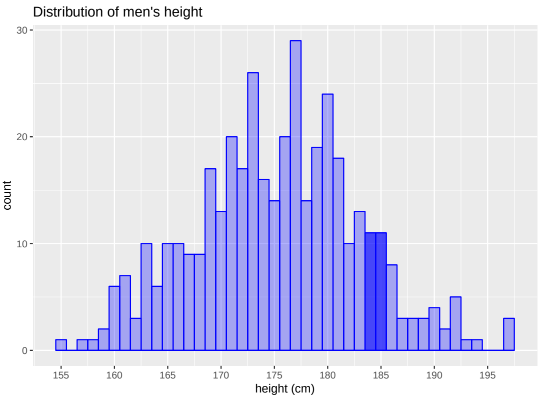


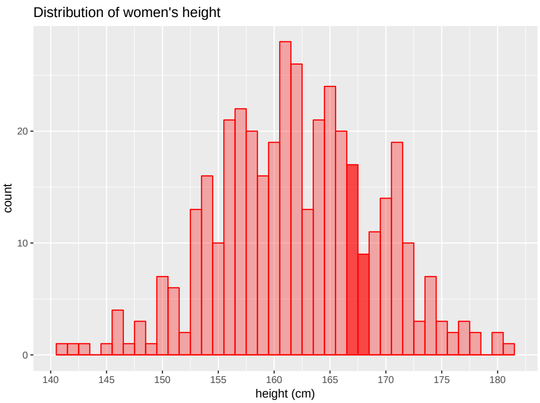
Normalizing Counts: Probability Density



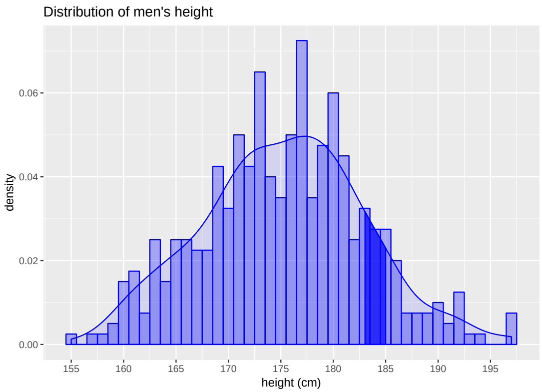


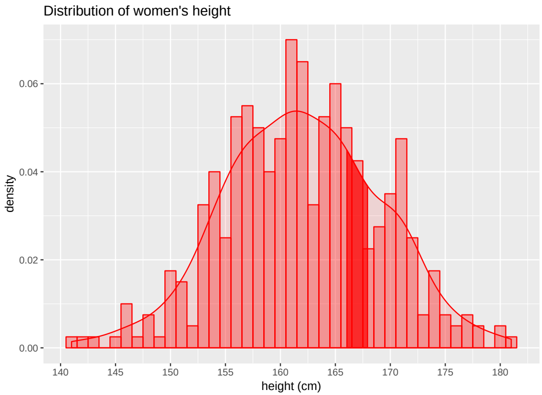
Computing Probability: Discrete Case





Computing Probability: Continuous Case

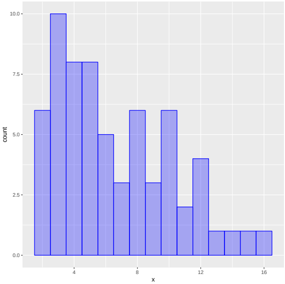




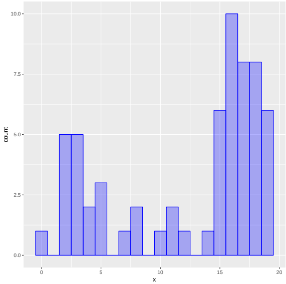
Common Distribution Shapes: Modality

Does the histogram have a single prominent peak? Two peaks? Several peaks? No obvious peaks?

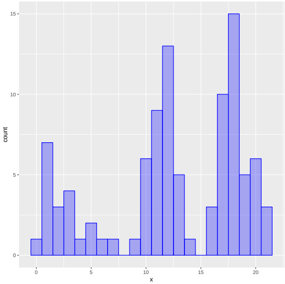
unimodal



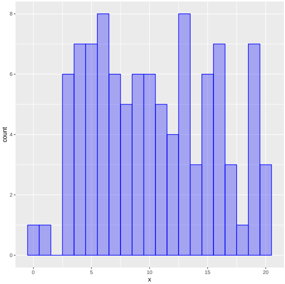
bimodal



multimodal



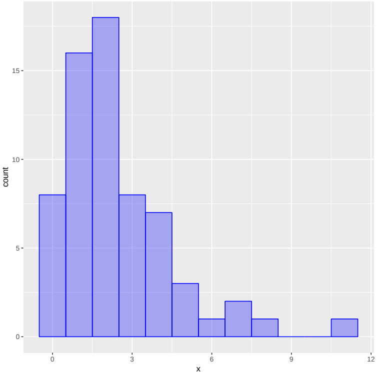
uniform



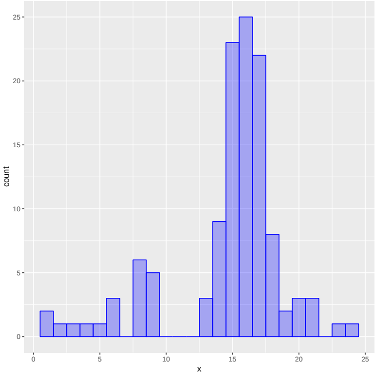
Imagining the corresponding density function may help.

Is the histogram *right skewed*, *left skewed*, or *symmetric*?

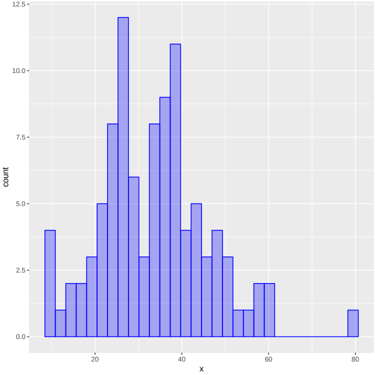
right skew



left skew



symmetric

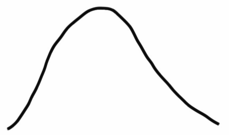


Histograms are said to be skewed to the side of the ***long tail.***

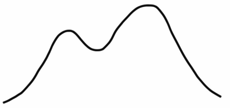
Common Distribution Shapes: Continuous

- modality

unimodal

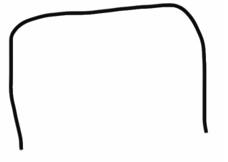


bimodal



multimodal

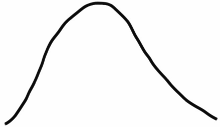
uniform



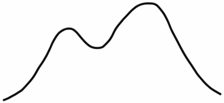
Common Distribution Shapes: Continuous

- modality

unimodal

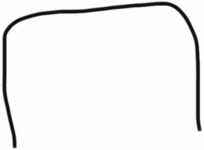


bimodal



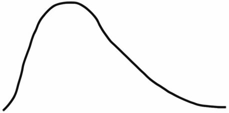
multimodal

uniform

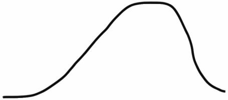


- skewness

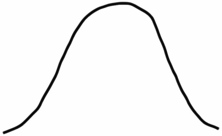
right skew



left skew

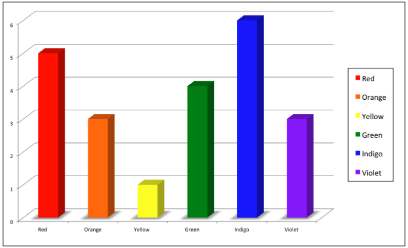


symmetric



Mode

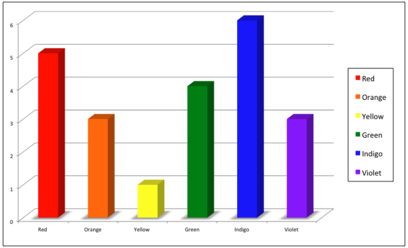
The most common, most probable, most popular, or most typical value of a random variable (outcome of a random process).



Central Tendency: Mode

Mode

The most common, most probable, most popular, or most typical value of a random variable (outcome of a random process).



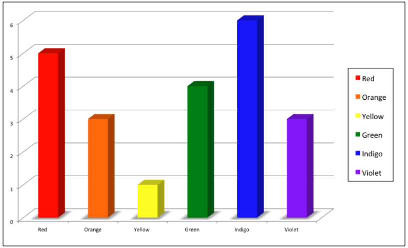
Less boring examples than “what is your favourite colour?”:

- *Which party will form a government after elections?*

Central Tendency: Mode

Mode

The most common, most probable, most popular, or most typical value of a random variable (outcome of a random process).



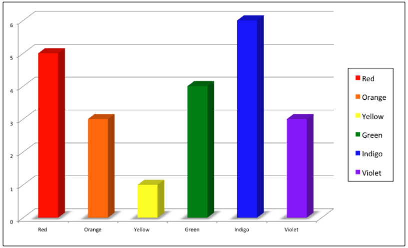
Less boring examples than "what is your favourite colour?":

- *Which party will form a government after elections?*
  The one that got the most votes (but this won't tell you how stable the coalition will be).

Central Tendency: Mode

Mode

The most common, most probable, most popular, or most typical value of a random variable (outcome of a random process).



Less boring examples than "what is your favourite colour?":

- *Which party will form a government after elections?*
  The one that got the most votes (but this won't tell you how stable the coalition will be).

NB: It does not make sense to compute the average colour or the average number of votes among the competing parties (except maybe to assess how fragmented the parliament is).

Central Tendency: Mean

Discrete case:
($p(x_i)$ is the probability of outcome x_i)

$$\bar{x} = \mu = \sum_i x_i p(x_i)$$

Continuous case:
($p(x)$ is the probability density function, PDF)

$$\bar{x} = \mu = \int_{-\infty}^{+\infty} x p(x) dx$$

If the probabilities of all outcomes x_i are the same (uniform distribution) then, for n possible outcomes, we have our usual definition of arithmetic mean:

$$p(x_i) = \frac{1}{n}, \qquad \bar{x} = \mu = \frac{1}{n} \sum_i x_i$$

If we know the probabilities of all the outcomes (we know the probability density function in the continuous case) then our mean is the *(mathematical) expectation*, *expected value*: $E[x] = \mu$.