- Random variables

- Probability mass/density functions

- Cumulative mass/density functions

- Expectation

A random variable is…

- a variable, and it is random…

A random variable is…

- a variable who takes on its value by chance. A random variable can take on a set of possible values, each with an associated probability.

- To characterise a random variable we need to know:
  - all possible outcomes (domain/support)
  - the probability of hitting each of the outcome (probability mass/density function)

- Let X be the outcome for tossing a fair coin.
  - X is a random variable
  - Two possible outcomes: $\{head, tail\}$
  - $\Pr(X = head) = 0.5, \Pr(X = tail) = 0.5$

- Let X be the outcome for rolling a fair die.
  - Six possible outcomes: $\{1, 2, 3, 4, 5, 6\}$
  - $\Pr(X = 1) = 1/6, \Pr(X = 2) = 1/6, \ldots$

- Let X be tomorrow's temperature.
  - Possible outcomes: from -15°C to 25°C.
  - Millions of possible outcomes. How can we quantify the probability then…?

Discrete and Continuous r.v.

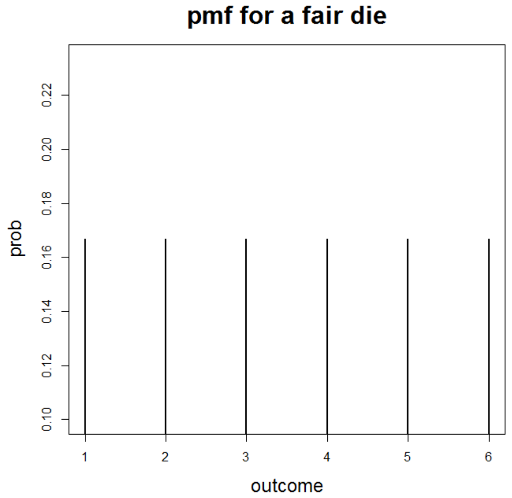
- A quantity X is called a discrete random variable if 1) it can only take a discrete collection of values, and 2) it is random.

- A quantity X is called a continuous random variable if 1) it can take a whole range of real-numbered values, and 2) it is random.

Probability mass function for discrete r.v.

- A probability mass function (or pmf) for a discrete random variable X is a function that describes the relative probability that X takes each of its possible values. Usually denoted by $f_X(x)$ or $f(x)$.

- If you plot it out, it looks like some vertical bars.



Probability density function for continuous r.v.

- A probability density function (or pdf) for a continuous random variable X is a function that describes the relative probability that X takes each value in the range of possible values.

- The range of possible values (with non-zero probability) is called the *support* of X.

Some common discrete r.v.

Bernoulli r.v.

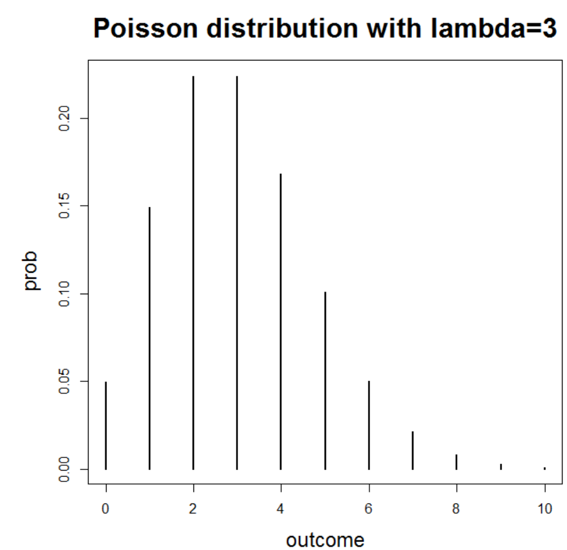
- Binary outcome: Success (1) or failure (0).
- One parameter: p. Probability of success.
- pmf: $f(X = 1) = p, f(X = 0) = 1 - p$
- Another expression: $f_X(x) = p^x(1 - p)^{1-x}$
- $X \sim Bernoulli(p)$

Binomial r.v.

- Sum of n independent and identically distributed Bernoulli r.v.
- Takes values on $\{0, 1, 2, \dots, n\}$
- Two parameters:
  - n. Number of independent Bernoulli trials.
  - p. Probability of success (inherited from Bernoulli r.v.).
- $f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$
- $X \sim bin(n, p)$

Poisson r.v.

- Number of events occurring in a fixed interval of time.
- Takes values on $\{0, 1, 2, \dots\}$.
- One parameter: $\lambda > 0$. Rate of events happening.
- $f(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
- $X \sim Poisson(\lambda)$



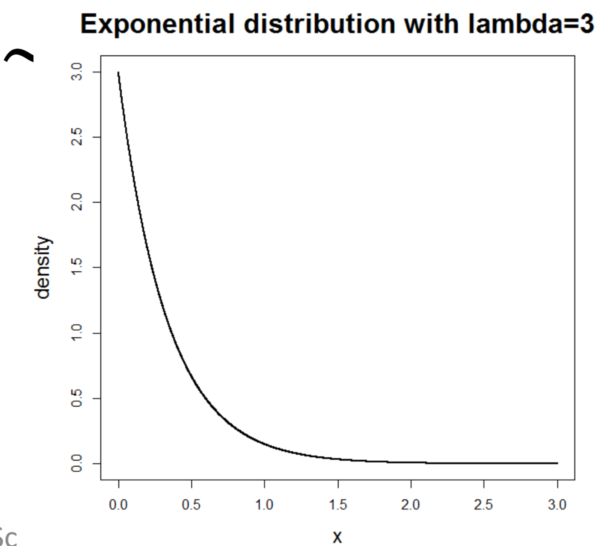
Some common continuous r.v.

Uniform r.v.

- Two parameters:
  - a: starting point
  - b: ending point
- $f_X(x) = \frac{1}{b-a}$
- $X \sim uniform(a, b)$

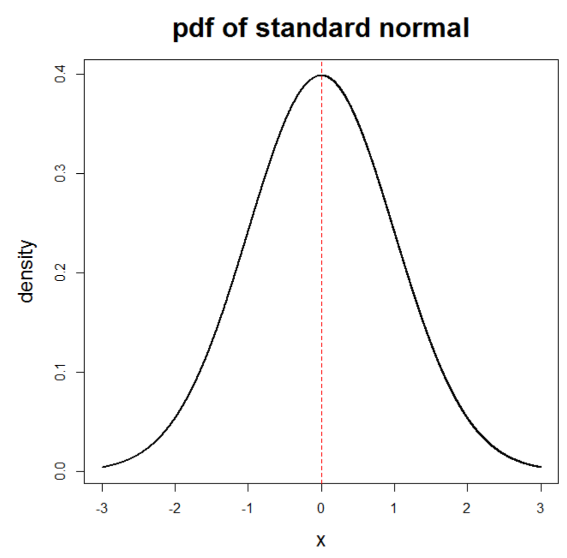
Exponential r.v.

- Time between events (remember Poisson?)
- Support: $[0, \infty)$
- λ: the rate parameter.
- $f_X(x) = \lambda \exp(-\lambda x)\,, x > 0$
- $X \sim \text{exponential}(\lambda)\,, or\ X\sim$



Normal r.v.

- The most famous one (we'll see why)
- Takes values over the real number line
- Two parameters: μ, σ^2
- $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$
- $X \sim N(\mu, \sigma^2)$



Properties of pmf/pdf

- Always above the x-axis (non-negative probability)
- [Discrete case] Sum of the probability mass (bars) =1
- [Continuous case] Area under the curve=1

- There are many more distributions:
  - Negative binomial, geometric, hypergeometric, gamma, beta, student’s t, chi-square, F, …
- Do think of what these r.v. represent. What kind of stochastic (random) processes are they referring to?

Example

$X \sim Poisson(\lambda)$ and its pmf is $f_X(k) = \frac{\lambda^k e^{-\lambda}}{k!}$.

Show that $\sum_{k=0}^{\infty} f_X(k) = 1$

Example

$X \sim \exp(\lambda)$ and its pdf is $f(x) = \lambda e^{-\lambda x}$. Show that $\int_{-\infty}^{\infty} f(x)dx = 1$

Cumulative mass/density function

- The probability of the r.v. X having the value less than of equal to x.
- $F(x) = \Pr(X \leq x)$, hence the name cumulative
- $F(-\infty) = 0\ and\ F(\infty) = 1$
- Always non-decreasing

- For discrete case, $F(x)$ is the sum of heights of the bars (from your pmf) with outcome values less than or equal to x.

- For continuous r.v., $F(x)$ is the area under the pdf curve, from negative infinity to x.

- $F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(t)dt$

- Show me your Calculus!

Expectation

Expectation

- If you repeat your experiment many many times (say, keep drawing some random numbers out, tossing a coin, throwing a die, …), then the expectation is the ‘average’ outcome you will get from these repeated experiments.

- Note that the ‘average’ here is a probabilistic statement (which is different from the average calculated from our sample).

Mean

- $E(X) = \sum_{all\ outcomes} xf(x)$

- $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$

- 'Average' value weighted according to the probability distribution.

- $E(X)$ is the population mean of the r.v. X. It is a measure of central tendency.

Variance

- $Var(X) = E(X^2) - [E(X)]^2$
  $= E[(X - E(X))^2]$

- Variance is a measure of dispersion

- $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$

Example

- $X \sim Bernoulli(p)$

$$E(X) = \sum xf(x)$$

$$= 0 * (1 - p) + 1 * p$$
$$= p$$

- “If you repeat your experiment for many times, independently, on average p of them will be a success.”

More on expectation

- $E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$,

  $for\ any\ real\ function\ g$

- $E(X+Y) = E(X) + E(Y)$ for any r.v. X and Y

- BTW, the sum of r.v. is another r.v.

Statistical moments

$E(X)$: central tendency, mean

$E(X^2)$: dispersion, variance

$E(X^3)$: skewness

$E(X^4)$: kurtosis

We called them moments. The n^{th} moment of a random variable is $E(X^n)$

Moment generating function

- The moment generating function (mgf) can also be used to characterise the distribution of r.v. Denoted by $M_X(t)$.
- The mgf 'generates' moments by the following way:
- $n^{th} moment = E(X^n) = \frac{d^n M_X(t)}{dt^n}|_{t=0}$