Probability Theory for EOR

How to describe the distribution

Distributions determine probability values of events generated by random variables. We distinguish random variables by their distributions.

How do we describe distributions: e.g., the shape of PDF/PMF?

Median and mode

Median.

- c **(may not be unique for a given r.v.)** is a median if $\mathbb{P}(X \leq c) \geq 1/2$ and $\mathbb{P}(X \geq c) \geq 1/2$.
- Median c is different from mean/expectation μ.
  - **The** value that minimizes the mean squared error, $\mathbb{E}(X-x)^2$, is **the mean** μ.
    Sketch of proof: $\mathbb{E}(X-x)^2 = \mathbb{E}(X-\mu+(\mu-x))^2 =$
    $\mathbb{E}(X-\mu)^2 + 2\mathbb{E}((X-\mu)(\mu-x)) + \mathbb{E}(\mu-x)^2 = \mathrm{Var}(X) + (\mu-x)^2$.
  - **A** value that minimizes the mean absolute error, $\mathbb{E}|X-x|$, is **a median** c.
    Sketch of proof: compare $E|X-x|$ with $\mathbb{E}|X-c|$ for $x<c$ and $x>c$.

Example:

- Median for Bern(p)-distributed X with $p=1/2$:
  $\forall c \in [0,1]$.
- Median for a constant c (degenerated random variable X, $P(X=c)=1$):
  c.

Mode.

- c **(may not be unique for a given r.v.)** is a mode if the PMF/PDF takes its maximum value at c.

Example:

- Mode for Bern(p)-distributed X with $p > 1/2$:

  $c = 1$. (Unimodal distribution, a probability distribution with a PDF/PMF which has a single peak).
- Mode for Bern(p)-distributed X with $p = 1/2$:

  $\{0, 1\}$. (Multimodal distribution, a probability distribution with a PDF/PMF which has multiple peaks, here we have a bimodal case).
- Mode for a constant c (degenerated random variable X, $P(X = c) = 1$):

  c.
- Mode for $N(\mu, \sigma^2)$-distributed X:

  $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ takes its maximum at $c = \mu$.

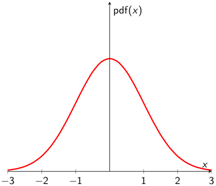
Symmetry and skewness

Symmetry.

- We say that an r.v. X has a symmetric distribution about μ (or X is symmetric) if $X-\mu$ has the same/identical distribution as $\mu - X$.

Example:

- If $Z \sim N(0,1)$, both Z and $-Z$ have the identical **PDF**:



$$\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

- Flip one coin, betting on the head is the same as betting on the bottom. Here we may consider $X = I_{head}$ (Bern(1/2)).
  Symmetric as **PMF** satisfies
  $P(X-\mu = k) = P(\mu - X = k)$.
  What else?
  $\mathbb{E}(X-\mu) \equiv 0$, $\mathbb{E}(X-\mu)^3 = 0$, $\mathbb{E}(X-\mu)^5 = 0, \cdots$.

Skewness.

- The skewness of an r.v. X with μ, σ^2 is:

$$\text{Skew}(X) = \mathbb{E}\left(\frac{X-\mu}{\sigma}\right)^3$$

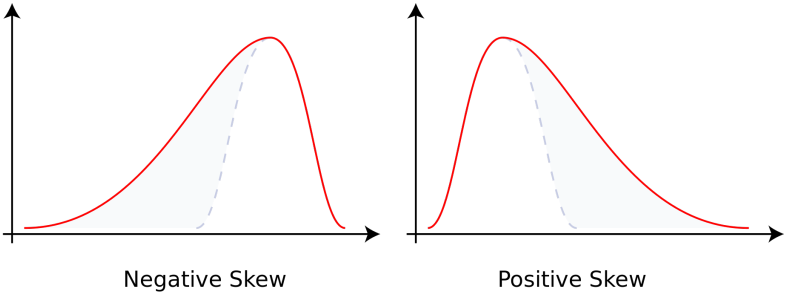
- We normalize by σ since a good measure of the symmetry should give the same number to measure the symmetry of X and $Y = bX$: **skewness is a third standardized moment**.

  Symmetry implies that $\text{Skew}(X) = 0$;
  Skew$(X) \neq 0$ **implies asymmetry.**

  Example:

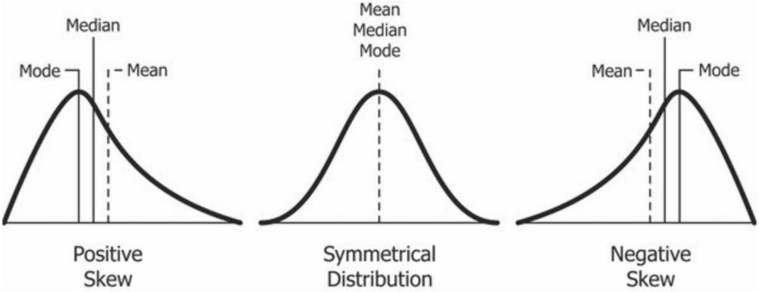
- Consider $Y \sim$Bern(1/2), $\mathbb{E}\left(\frac{Y-1/2}{1/2}\right)^3 = -1 \times 1/2 + 1 \times 1/2 = 0.$
- **Skew**$(X) = 0$ **does not necessarily imply symmetry**:
  Consider a discrete r.v. X such that supp(X)={-1,-2,3} with **PMF** $p_X(x) = 1/3, x \in \{-1, -2, 3\}$, X is not symmetric.

Example figure I:
Figures from wikipedia (https://en.wikipedia.org/wiki/Skewness) on negative (left) skewed and positive (right) skewed distributions.



Example figure II:

Figures from wikipedia (https://en.wikipedia.org/wiki/Skewness) on negative (left) skewed and positive (right) skewed **unimodal** distributions.



Tail and Kurtosis

Kurtosis.

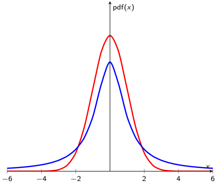
- The Kurtosis of an r.v. X with μ, σ^2 is (**fourth standardized moment shifted by 3**):

$$\mathrm{Kurt}(X) = \mathbb{E}\left(\frac{X-\mu}{\sigma}\right)^4 - 3$$

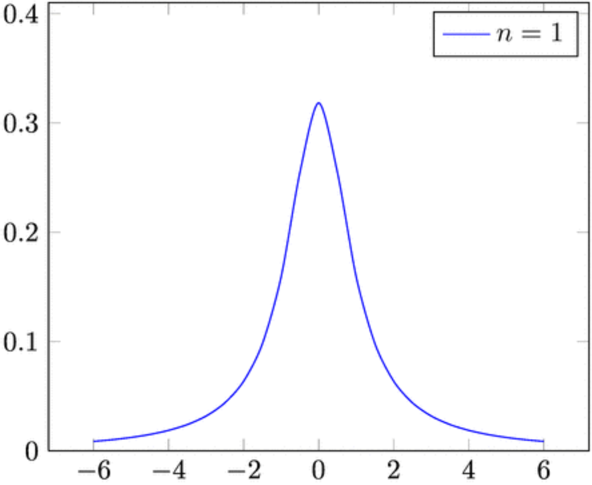
- A measure of how heavy the tail of the probability distribution of a real-valued random variable is.

Example:

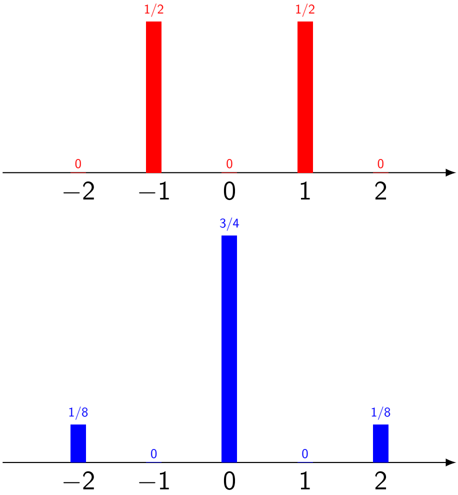
- Student-t(n) PDF (n=5) (Kurtosis is 5) and standard normal bell-shape PDF (Kurtosis is 0)



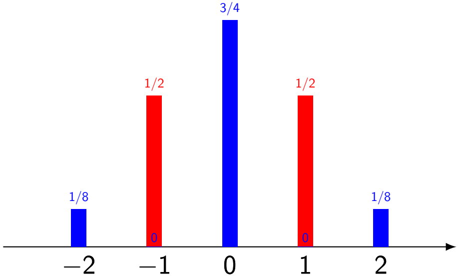
Student-t(n) PDF:



Example:
Fair coin v.s. unfair 3-sided die



Example:
Fair coin v.s. unfair 3-sided die



1. Mean the same
2. Variance the same
3. Symmetric (Skew zero)
4. ** Kurt(red) < Kurt(blue)