Chapter 4 sections

- 4.1 Expectation
- 4.2 Properties of Expectations
- 4.3 Variance
- 4.4 Moments
- 4.5 The Mean and the Median
- 4.6 Covariance and Correlation
- 4.7 Conditional Expectation
- SKIP: 4.8 Utility

Median – measure of center

Def: Median

Let X be a random variable. Every number m that satisfies

$$P(X \leq m) \geq 0.5 \quad \text{and} \quad P(X \geq m) \geq 0.5$$

is called a *median* of the distribution of X.

- Recall the distribution of $Y =$ the number of heads in 3 tosses (coin toss example from Lecture 4)

y	0	1	2	3
$f_Y(y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Both 1 and 2 are medians since

$$P(X \leq 1) = 1/2 \quad \text{and} \quad P(X \geq 1) = 7/8 \geq 0.5$$
$$\text{and} \quad P(X \leq 2) = 7/8 \geq 0.5 \quad \text{and} \quad P(X \geq 2) = 1/2$$

In fact all numbers in the interval $[1, 2]$ are medians.

Example 1: Median of the exponential distribution

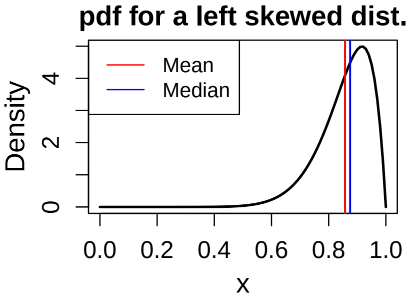
Let $X \sim \mathrm{Expo}(\lambda)$. The pdf of X is

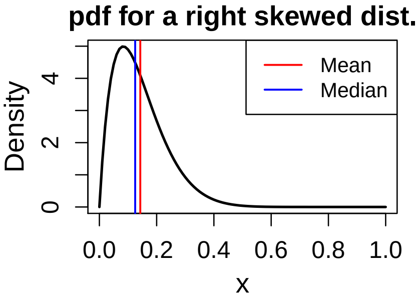
$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

Find the median of X

Median – measure of center

- For a symmetric distribution: mean = median
- For a skewed distribution either
  - mean < median (left skewed) or mean > median (right skewed)





Mean vs. median

Both mean and median can be used as a mesure of center

- Median exists for every distribution but the mean may not exist
- For a skewed distribution the mean is heavily influenced by the tail ("outliers") but the median is not

Mean and Median can be used as estimates of a random variable

- The mean minimizes the mean squared error (Theorem 4.5.2)
- The median minimized the mean absolute error (Theorem 4.5.3)

Covariance

Def: Covariance

Let X and Y be random variables with finite means μ_X and μ_Y. The *covariance of X and Y* is defined as

$$\mathrm{Cov}(X, Y) = E\left((X - \mu_X)(Y - \mu_Y)\right)$$

if the expectation exists.

- Another way of calculating the covariance:

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

- A measure of how much X and Y depend (linearly) on each other
- Magnitude of $\mathrm{Cov}(X, Y)$ depends on the magnitudes of X and Y

Correlation

Def: Correlation

Let X and Y be random variables with finite variances σ_X^2 and σ_Y^2. The *correlation of X and Y* is defined as

$$\mathrm{Cor}(X,Y) = \rho(X,Y) = \frac{\mathrm{Cor}(X,Y)}{\sigma_X \sigma_Y}$$

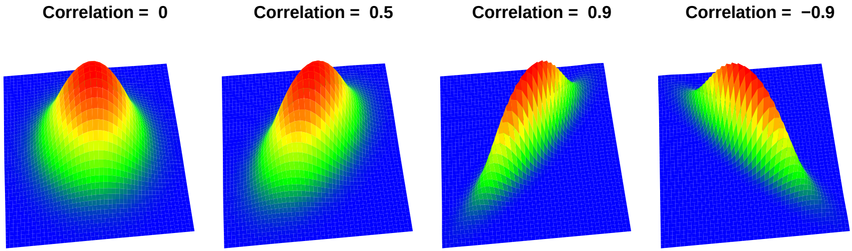
- $-1 \leq \rho(X,Y) \leq 1$
  - This is shown by using the *Schwarz Inequality*:

$$(E(UV))^2 \leq E(U^2)E(V^2)$$

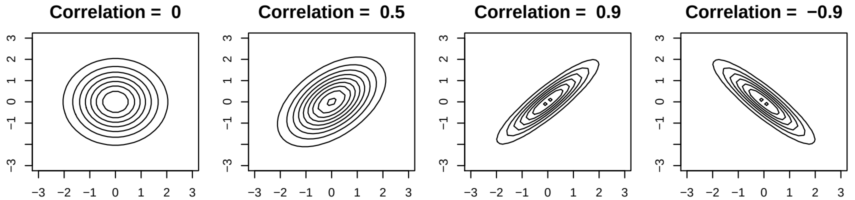
- Also a measure of how much X and Y depend (linearly) on each other
- But the correlation is independent on the scale of X and Y

Example 2: Bivariate Normal

Joint pdf of two correlated Gaussian random variables:



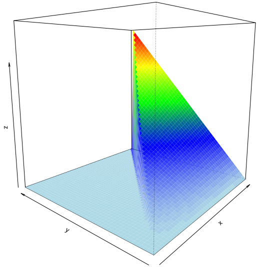
Contours of the joint pdf of two correlated Gaussian random variables:



Example 3: Calculating Covariance and Correlation

Recall from Lecture 4 the joint pdf for random variables X and Y:

$$f(x,y) = 8xy \quad \text{for } 0 < y < x < 1$$



We have already established that the marginal pdf's are

$$f_X(x) = 4x^3 \quad \text{for } 0 < x < 1$$
$$f_Y(y) = 4y - 4y^3 \quad \text{for } 0 < y < 1$$

Find $\mathrm{Cov}(X,Y)$ and $\rho(X,Y)$.

Properties of covariance and correlation

Theorem

If r.v. X and Y have finite variances, i.e. $\sigma_X^2 < \infty$ and $\sigma_Y^2 < \infty$, then the covariance $\text{Cov}(X, Y)$ exists.

Theorem 4.6.4

If X and Y are independent random variables with finite variances then

$$\text{Cov}(X, Y) = \rho(X, Y) = 0$$

- Careful: The oposite is not true, i.e. two random variables can be uncorrelated without being independent.

Example 4: Zero covariance does not imply independence

- Let X and Z be independent random variables where $X \sim \mathrm{Uniform}(-1, 1)$ and $Z \sim \mathrm{Uniform}(0, 0.1)$.
- Let $Y = X^2 + Z$. Then Y and X are clearly not independent.
- Show that $\mathrm{Cov}(X, Y) = 0$

Properties of covariance and correlation

Other properties:

- If Y is a linear function of X then X and Y are perfectly correlated, i.e. $\rho(X, Y) = \pm 1$
- $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
- $\mathrm{Cov}(aX + b, cY + d) = ac\mathrm{Cov}(X, Y)$ (Homework, Ex. 4.6.5)
- $\mathrm{Var}(aX + bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y)$

  Note that the Cov part is zero if X and Y are independent.

Conditional Expectation

Def: Conditional Mean

Let X and Y be random variables where Y has a finite mean. The *conditional expectation of Y given $X = x$*, denoted $E(Y|x)$ or $E(Y|X = x)$, is the mean of the conditional distribution of Y given $X = x$

More explicitly:

$$E(Y|x) = \int_{-\infty}^{\infty} yf(y|x)dy \quad \text{if } Y \text{ is continuous}$$

$$E(Y|x) = \sum_{\text{All } y} yf(y|x) \qquad \text{if } Y \text{ is discrete}$$

- $E(Y|x)$ is a function of x
- $E(Y|X)$ is a random variable
  (i.e. a function of the random variable X)

Conditional Variance

Def: Conditional Variance

Let X and Y be random variables. The *conditional variance of Y given $X = x$*, denoted $Var(Y|x)$ or $Var(Y|X = x)$, is the variance of the conditional distribution of Y given $X = x$:

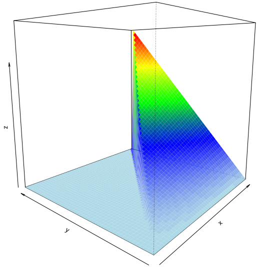
$$\mathrm{Var}(Y|x) = E\left((Y - E(Y|x))^2 \middle| x \right)$$

- $\mathrm{Var}(Y|x)$ is a function of x
- $\mathrm{Var}(Y|X)$ is a random variable

Example 5: Calculating conditional mean and variance

Consider again the joint pdf for random variables X and Y:

$$f(x,y) = 8xy \quad \text{for } 0 < y < x < 1$$



We have already established that the marginal pdf's are

$$f_X(x) = 4x^3 \quad \text{for } 0 < x < 1$$
$$f_Y(y) = 4y - 4y^3 \quad \text{for } 0 < y < 1$$

Find $E(Y|X = 0.5)$ and $\mathrm{Var}(Y|X = 0.5)$

Law of total probability for E and Var

Theorem 4.7.1: Law of total probability for Expectations

Let X and Y be random variables such that Y has finite mean. Then

$$E\left(E(Y|X)\right) = E(Y)$$

Theorem 4.7.4: Law of total probability for Variances

Let X and Y be random variables. Then

$$\mathrm{Var}(Y) = E\left(\mathrm{Var}(Y|X)\right) + \mathrm{Var}\left(E(Y|X)\right)$$

Example: Hierarchical Model

Using laws of total probability

Screening for defective Halloween candy among n pieces

1. My daughter screens all the n items, the probability that an item passes her screening is p_X.
2. I screen what passed my daughters screening and the probability that a candy passes my screening is p_Y

How many candies can we expect pass this double screening?

- X = number of candies that pass first screening.
  $X \sim \mathrm{Binomial}(n, p_X)$
- Y = number of candies that pass second screening.
  $Y|X = x \sim \mathrm{Binomial}(x, p_Y)$

Find $E(Y)$ and $\mathrm{Var}(Y)$

END OF CHAPTER 4