Chapter 4: Expected Values

In this chapter we will cover:

1. The expected value of a random variable , (§4.1 Rice)
2. Variance and standard deviation, (§4.2 Rice)
3. Covariance and correlation (§4.3 Rice)

□

The expected value

- If X is a discrete random variable with frequency function $p(x)$ then the expected value, denoted by $E(X)$ is defined as
$$E(X) \sum_i x_i p(x_i)$$
when this sum 'makes sense'
- $E(X)$ is also called the mean of X and denoted often by μ

□

Roulette

- A roulette wheel has numbers 1 to 36 and 0 and 00
- If you bet on \$1 on an odd number you win (or lose) \$1 depending on if it comes up.
- Let X be yours winnings, this has a probability frequency function
$$p(x) = \begin{cases} \frac{18}{38} & x = 1 \\ \frac{20}{38} & x = -1 \end{cases}$$
- So the expected value is
$$1 \times \frac{18}{38} + (-1) \times \frac{20}{38} = -\frac{1}{19}$$

□

Poisson random variables

- If X is a poisson random variable its frequency function is $\frac{\lambda^k}{k!} \exp(-\lambda)$
- Its expectation is then
$$\begin{aligned} E(X) &= \sum_{i=0}^{\infty} k \times \frac{\lambda^k}{k!} \exp(-\lambda) \\ &= \lambda \exp(-\lambda) \sum_{i=1}^{\infty} k \times \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda \exp(-\lambda) \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\ &= \lambda \end{aligned}$$

□

Continuous random variables

- If X is a continous random variable with density function $f(x)$ then the expected value is defined as

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

when this integral 'makes sense'

- $E(X)$ is also called the mean of X and denoted often by μ

□

Normal distribution

- Suppose that X has a Normal (μ, σ^2) distribution
- The expected value of X is then

$$\int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\; dx$$

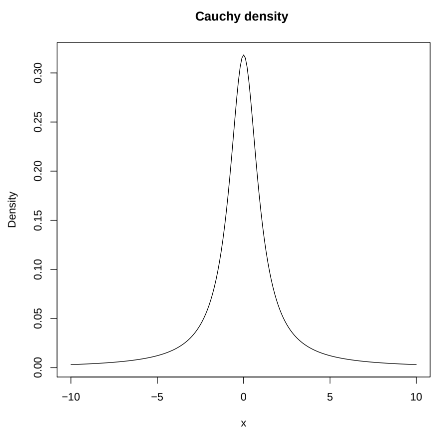
- This equals, see page 114 Rice, μ.

□

Cauchy distribution

- The cauchy density is

$$\frac{1}{\pi}\frac{1}{1+x^2}$$



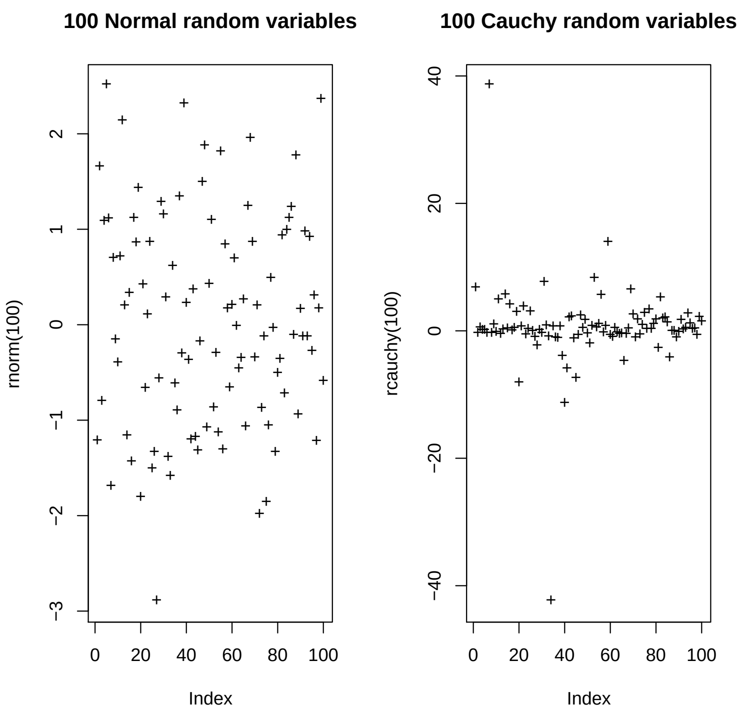
- The integral

$$\int_{-\infty}^{\infty} x \frac{1}{\pi}\frac{1}{1+x^2}\; dx$$

does not exist (i.e., is infinite)

□

Cauchy distribution



Expectations of functions of R.V.

- Suppose that $Y = g(X)$
- If X is discrete then

$$E(Y) = \sum_x g(x)p(x),$$

  when this sum is defined
- If X is continuous then

$$E(Y) = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

  when this integral is defined
- Note that in general

$$E(g(X)) \neq g(E(X))$$

  for most functions g.

Expectations of linear functions of R.V.

- Let $X_1, X_2, \cdots, X_n$ be a set of random variables with expectations $E(X_i)$
- Let $Y = a + \sum_{i=1}^n b_i X_i$ be a linear conbination of the X_i's
- Then

$$E(Y) = a + \sum_{i=1}^{n} b_i E(X_i)$$

See Thoerem A page 119 Rice.

□

Example

- Suppose that Y has a binomial distribution
- Directly its expectation is

$$\sum_{i=1}^{n} k \times \binom{n}{k} p^k (1-p)^{n-k}$$

which is not easy to evaluate

- Since Y is binomial it is the sum of n Bernoulli r.v. X_i, each of these has $E(X_i) = 1 \times p + 0 \times (1-p) = p$
- Hence using the previous result gives $E(Y) = np$

□

Exercise

1. Suppose that X is a discrete uniform random variable i.e. $P(X = k) = 1/n$ for $i = 1, 2, \cdots, n$. (a) What is $E(X)$? (b) What is $E(X^2)$?
2. Suppose that X is a continuous random variable with density $f(x) = 2x$ for $0 \leq x \leq 1$ (and 0 outside this interval). What is (a) $E(X)$? (b) What is $E(X^2)$?

□

Recommended Questions

From §4.7 Rice Questions 4(a), 7(a, b), 8, 13, 27 (see ex A,B p.120)

□

Variance and standard deviation

- The expect value of a random variable is an average quantity and can be thought of as the 'central value' of the density or frequency function
- It is often called a **location** parameter
- The standard deviation is another parameter which indicated how dispersed the probability density function is
- First define the *variance* of a random variable by

$$\mathrm{Var}(X) = E\{[X - E(X)]^2\}$$

  assuming all expectations exist.
- The standard deviation is the square root of the variance

□

Variance and standard deviation

- If X is discrete with frequency function $p(x)$

$$\mathrm{Var}(X) = \sum_i (x_i - \mu)^2 p(x_i)$$

  assuming the sums make sense.
- If X is continuous with density function $f(x)$ then

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx$$

- The variance is often denoted by σ^2 and the standard deviation by σ

□

Variance and standard deviation

- If $Y = a + bX$ then

$$\mathrm{Var}(Y) = b^2 \mathrm{Var}(X)$$

  see page 123, Theorem A

□

Bernoulli distribution

- The mean for a Bernoulli random variable X is

$$E(X) = 1 \times p + 0 \times (1 - p) = p$$

- The variance is given by

$$\begin{aligned}\mathrm{Var}(X) &= (1-p)^2 \times p + (0-p)^2 \times (1-p) \\ &= p - 2p^2 + p^3 + p^2 - p^3 \\ &= p(1-p)\end{aligned}$$

□

Variance

- An alternative way of calculating the variance is given by the following result

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

- For a uniform $[0,1]$ random variable X we have $E(X) = \frac{1}{2}$
- $E(X^2) = \int_0^1 x^2\, dx = \frac{1}{3}$
- So $\text{Var}(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$

□

Chebyshev's inequality

- Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}$$

- This can be rewritten as for all $k > 0$

$$P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}$$

- That is any random varibles will be very unlikely to be a large number of standard errors from the mean

□

Recommended questions

From Rice §4.7 look at questions: 4(a), 7(c,d), 14, 15, 16, 33

□

Covariance and Correlation

- Up to now our random variables have been single numbers
- This can be generalised to a *joint* distribution, where we have two numbers X and Y whose values are determined by chance
- If they are continous then there exists a joint density function $f(x, y)$ such that for any subset $A \subset \mathbf{R} \times \mathbf{R}$

$$P(\omega \in A) = \int_A f(x, y)\, dxdy$$

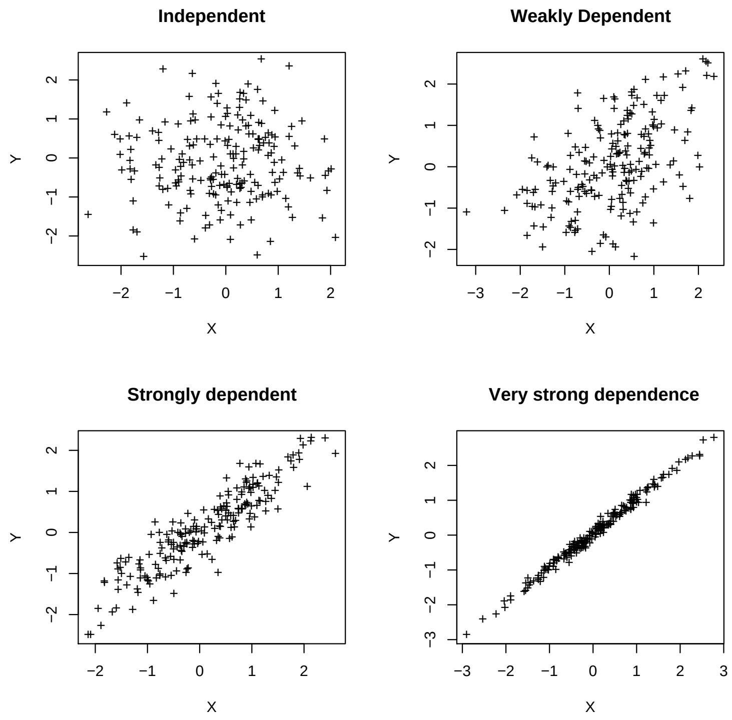
  where now we have a double integral

- If X and Y are independent then $f(x, y) = f_X(x) f_Y(y)$

□

Dependence

The figure shows types of dependence for jointly distributed random variables



In each case what does learning about X tell us about Y?

Covariance and Correlation

- If X and Y are jointly distributed random variables with expectations μ_X and μ_Y then the covariance of X and Y is

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

- This can be written, see page 130, as

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

- If X and Y are independent then $E(XY) = E(X)E(Y)$ so there covariance will be zero.

Linear Transformations

1. If $U = X + a$ where a is a constant then

$$\mathrm{Cov}(U, Y) = \mathrm{Cov}(X, Y)$$

2. If $U = aX$ then

$$\mathrm{Cov}(U, Y) = a\mathrm{Cov}(X, Y)$$

3. If X, Y and Z are three random variables then

$$\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$$

4. For all X it is clear from the definition that

$$Cov(X, X) = \mathrm{Var}(X)$$

5. For any X, Y

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$

□

Variance of Binomial

- The bimomial random variable is the sum of in independent number of Bernoulli's
- The variance of a Bernoulli is $p(1 - p)$
- Hence the variance of a binomial is $np(1 - p)$

□

Correlation

- The correlation of two random variables X and Y is given by

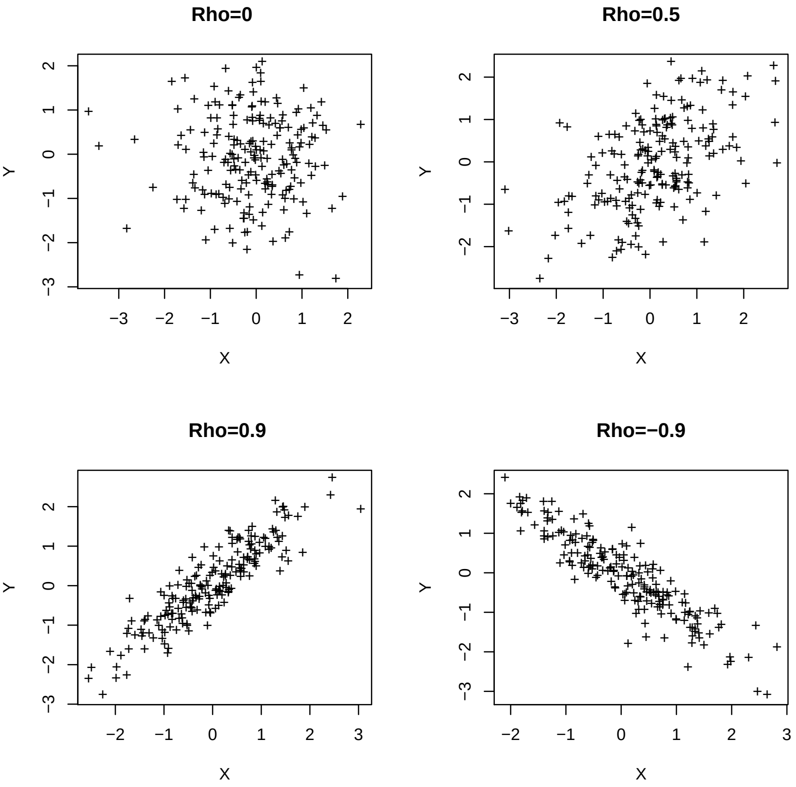
$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

- Note that this (unlike covariance) is a dimensional less quantity. So its does not matter what units X and Y are measured in.

□

Correlation

The figure shows different correlations



□

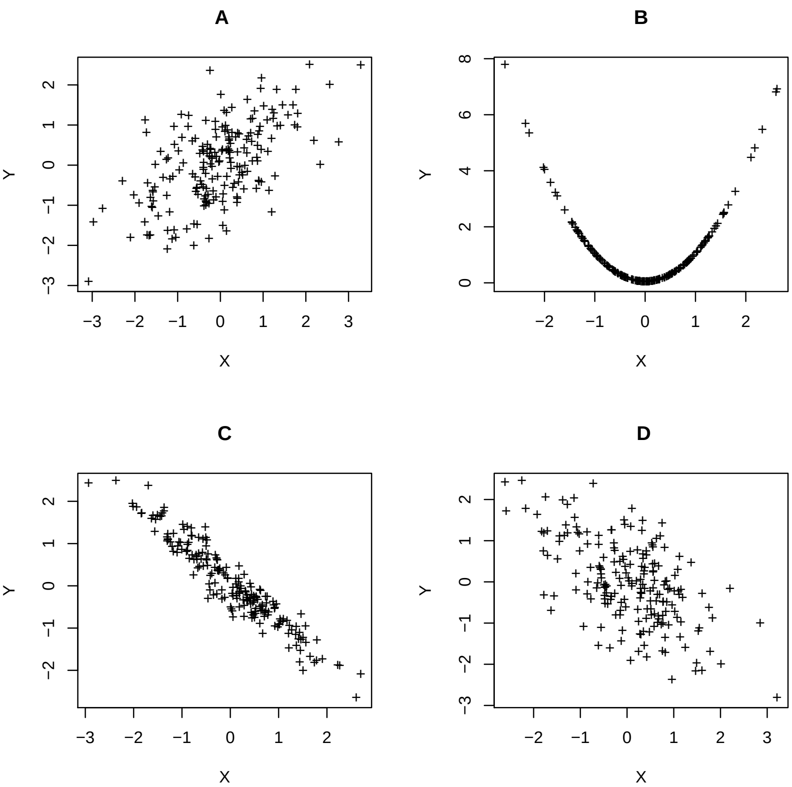
Correlation

- From Theorem B, page 133 we have that
$$-1 \leq \rho \leq +1$$
- If $\rho = \pm 1$ if and only if $Y = a + bX$ for some constants a, b.
- We say that the correlation measure the strength of the *linear* relationship between X and Y
- It does not directly measure dependence

□

Exercise

Match the following correlations to the plots: $0.6, -0.6, -0.95, 0$



Which plots show the strongest dependence? □

Recommended questions

From Rice §4.7 look at questions: 39, 40, 42,43, 49 □