Chapter 4.1 Random Variables

A **random variable** is a variable that takes on numerical outcomes defined over a sample space of a random experiment.
A random variable has a probability distribution.

A random variable can be denoted by **X** (upper-case) and a possible numerical outcome is x (lower-case).

Example: The random variable **X** is age in years of a UBC student. The possible outcomes are:

x = 15, 16, 17, 18, . . . etc.

Types of random variables:

- A **discrete random variable** has a countable number of values (typically integer numbers). **limited number of outcomes**

  Example 1: age in years of a UBC student

  Example 2: categorical variables.
  For example, the random variable **X** represents gender.
  The possible values can be assigned the codes:

  x = 0 male

  x = 1 female

- A **continuous random variable** can take any numerical value in an interval of the real number line.
  Examples: income, stock market prices, interest rates, consumer price index, etc.

Chapter 4.2 Discrete Random Variables

For a discrete random variable **X** the **probability distribution function** is:

$$\mathbf{P(x) = P(X = x)} \qquad \text{for all possible values of x.}$$

Example: The random variable **X** is the number resulting from the throw of a six-sided dice.
The probability distribution function is:

x	1	2	3	4	5	6
P(x)	$\mathbf{1/6}$	$\mathbf{1/6}$	$\mathbf{1/6}$	$\mathbf{1/6}$	$\mathbf{1/6}$	$\mathbf{1/6}$

That is, $\mathbf{P(x) = 1/6}$ for x = 1, 2, 3, 4, 5, 6

The probability distribution function for a discrete random variable has the properties:

- $\mathbf{0 \le P(x) \le 1}$ for all possible values of x.
- $\mathbf{\sum_{x} P(x) = 1}$

  ↑

  summation over all possible values of x.

The **cumulative probability function** is defined as:

$$\mathbf{F(a) = P(X \leq a)} \qquad \text{for all possible values of } \mathbf{a}.$$

This can be calculated from the probability distribution function as:

$$\mathbf{F(a) = \sum_{x \leq a} P(x)}$$

↑

summation over all possible values of x that are less than or equal to **a**.

Example: For the dice throwing experiment:

$$\mathbf{F(1) = P(X \leq 1) = P(X = 1) = 1/6}$$

$$\mathbf{F(2) = P(X \leq 2) \quad = P(X = 1) + P(X = 2)}$$

$$\mathbf{= 1/6 + 1/6 = 1/3}$$

$$\mathbf{F(3) = P(X \leq 3) \quad = P(X = 1) + P(X = 2) + P(X = 3)}$$

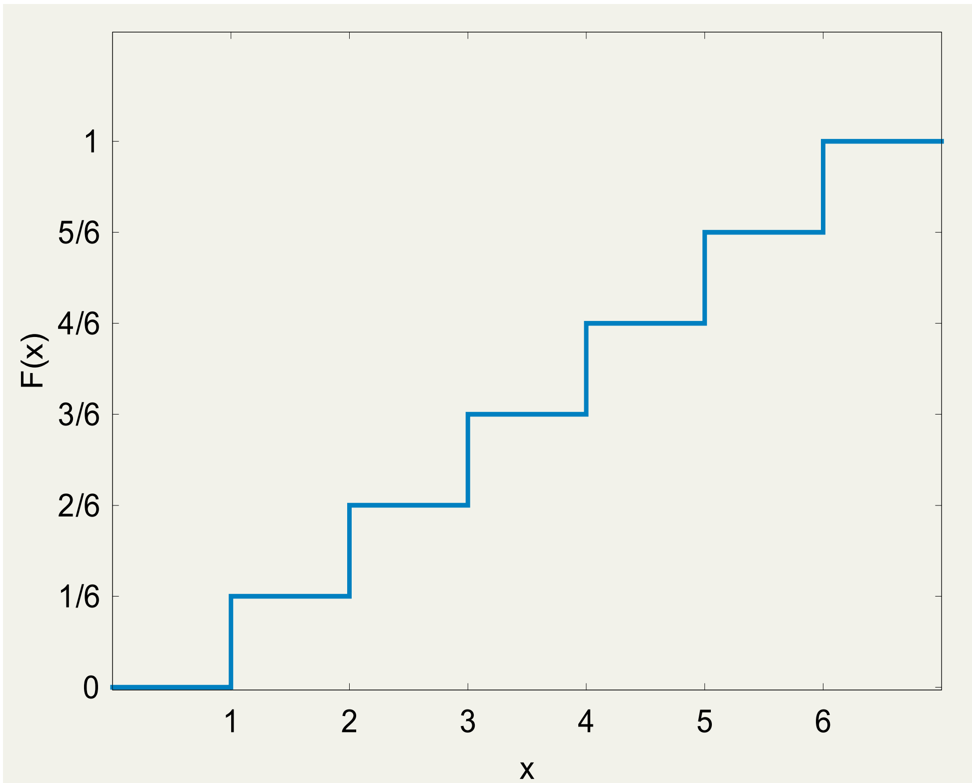
$$\mathbf{= 1/2}$$

$$\mathbf{F(4) = P(X \leq 4) = 2/3}$$

$$\mathbf{F(5) = P(X \leq 5) = 5/6}$$

$$\mathbf{F(6) = P(X \leq 6) = 1}$$

Graph of the cumulative probability function for
the dice throwing experiment.



The graph illustrates that, for a discrete random variable, the cumulative probability function is a step function that begins at 0 and ends at 1.

The cumulative probability function for a discrete random variable has the properties:

- $0 \le F(a) \le 1$ for all possible values of **a**.
- For two numbers **a**, **b** with $a < b$ then

  $$F(a) \le F(b)$$

- $$P(X > a) = 1 - P(X \le a)$$
  $$= 1 - F(a)$$

Example

The random variable **X** is the number of flights delayed per hour at an international airport.
The probability distribution function and cumulative probability function are:

x	P(x)	F(x)
0	0.10	0.10
1	0.08	0.18
2	0.07	0.25
3	0.15	0.40
4	0.12	0.52
5	0.08	0.60
6	0.10	0.70
7	0.12	0.82
8	0.08	0.90
9	0.10	1.00

What is the probability of five or more delayed flights in a given hour ?

$$
\begin{aligned}
\mathbf{P(X \geq 5)} &= \mathbf{1 - P(X \leq 4)} \\
&= \mathbf{1 - F(4)} \\
&= \mathbf{1 - 0.52} \\
&= \mathbf{0.48}
\end{aligned}
$$

What is the probability of three through seven (inclusive) delayed flights in a given hour ?

$$
\begin{aligned}
\mathbf{P(3 \leq X \leq 7)} &= \mathbf{P(X \leq 7) - P(X \leq 2)} \\
&= \mathbf{F(7) - F(2)} \\
&= \mathbf{0.82 - 0.25} \\
&= \mathbf{0.57}
\end{aligned}
$$

Mean = expected value bei discrete random variables

Chapter 4.3 Mean and Variance of Discrete Random Variables

Summary measures of the information in the probability distribution are of interest. Recall, the mean is the measure of central location for a data set of numeric observations. For a random variable, the **expected value** is the corresponding measure of central location.

For a discrete random variable **X**, the expected value is defined as:

$$\mathbf{E(X) = \sum_{x} x\,P(x)}$$

Erwartungswert = Wahrscheinlichkeit * Ergebnis

This is a weighted average of all possible outcomes where the weights are the probabilities.

The expected value of a random variable is also called its **mean** and is denoted by:

$$\mathbf{\mu_X = E(X)}$$

↑ (X: optional)

Greek letter mu

Example

The random variable **X** is the sum of the two numbers shown on a throw of two dice.

The probability distribution function of **X** is:

x	2	3	4	5	6	7	8	9	10	11	12
P(x)	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

The expected value is calculated as:

$$\mathbf{E(X)} = 2\left(\frac{1}{36}\right) + 3\left(\frac{2}{36}\right) + 4\left(\frac{3}{36}\right) + \ldots + 11\left(\frac{2}{36}\right) + 12\left(\frac{1}{36}\right)$$

$$= \mathbf{7}$$

P(X = 2) = P(dice 1 = 1) · P(dice 2 = 1) → independence

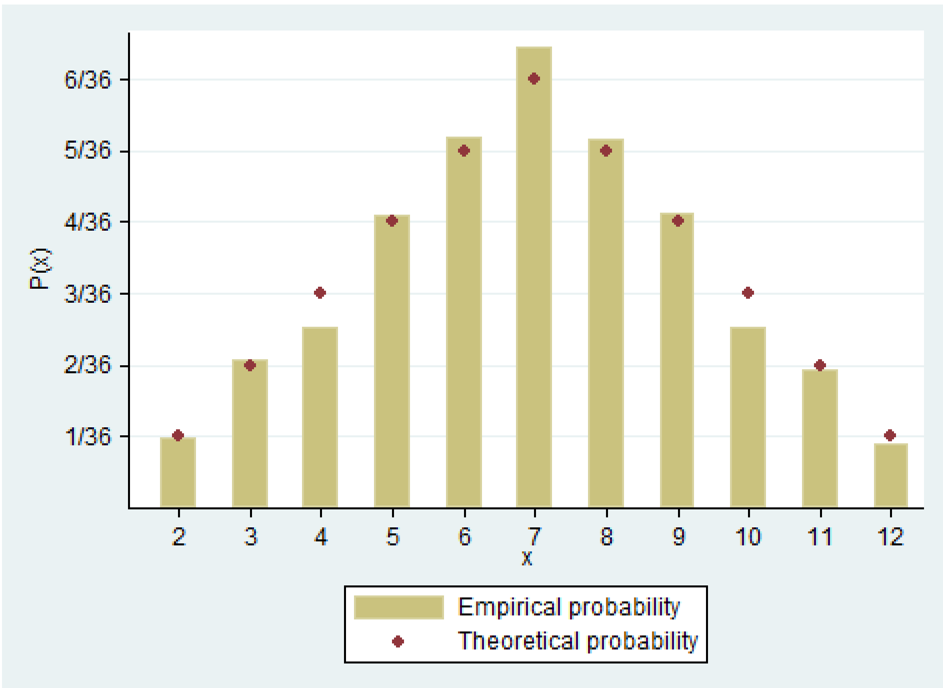
P(X = 3) = P(dice 1 is 2, dice 2 is 1) = $2 \cdot \left(\frac{1}{6}\right) \cdot \left(\frac{1}{6}\right)$

or P(dice 1 is 1, dice 2 is 2)

The expected value can be viewed as the long-run average value that a random variable would take over a "large" number of trials of the random experiment.

This can be illustrated for the dice-toss experiment.
A computer program was used to simulate the throw of two dice.
At each throw, the sum of the two dice faces was recorded.
This was repeated 2000 times.

The graph shows a histogram of the relative frequencies for the dice outcomes. This gives the empirical probability distribution function.



The average of the 2000 observed outcomes was 6.969, close to the expected value of 7.0.

Suppose $\mathbf{g(X)}$ is some function of the random variable $\mathbf{X}$. Then:

$$\mathbf{E[g(X)] = \sum_{x} g(x)\,P(x)}$$

Example: For $\mathbf{g(X) = X^2}$ the expectation is calculated as:

$$\mathbf{E(X^2) = \sum_{x} x^2\,P(x)}$$

A measure of dispersion for a random variable $\mathbf{X}$ is the **variance** defined as:

$$\mathbf{Var(X) \quad = E[\,(X - \mu_X\,)^2\,]}$$

Erwartungswert*
Abstand vom mean²

where $\mathbf{\mu_X = E(X)}$

$$\mathbf{= \sum_{x} (x - \mu_X)^2\,P(x)}$$

The variance of a random variable $\mathbf{X}$ is denoted by the symbol $\boldsymbol{\sigma_X^2}$ (sigma-squared):

$$\boldsymbol{\sigma_X^2} = \mathbf{Var(X)}$$

The variance can be expressed in an alternative way:

$$
\begin{aligned}
\mathbf{Var(X)} &= \sum_x (x - \mu_X)^2 P(x) \\
&= \sum_x (x^2 - 2\mu_X x + \mu_X^2) P(x) \\
&= \sum_x x^2 P(x) - 2\mu_X \sum_x x P(x) + \mu_X^2 \sum_x P(x) \\
&= \sum_x x^2 P(x) - 2\mu_X \cdot \mu_X + \mu_X^2 (1) \\
&= \sum_x x^2 P(x) - \mu_X^2 \\
&= E(X^2) - \mu_X^2
\end{aligned}
$$

That is, a useful calculation formula for variance is:

$$\mathbf{Var(X)} = E(X^2) - \mu_X^2 = \sum_x x^2 P(x) - \mu_X^2$$

The **standard deviation** of a random variable **X** is defined as:

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{\mathbf{Var(X)}} > 0$$

Note: A positive variance, and, therefore, positive standard deviation, assumes at least two distinct outcomes.

Consider two random variables **X** and **Y** with means:

$$\mu_X = E(X), \qquad \mu_Y = E(Y)$$

and variances:

$$\sigma_X^2 = Var(X), \qquad \sigma_Y^2 = Var(Y)$$

The two random variables may have the same mean but substantial differences in the variance. Suppose that:

$$\mu_X = \mu_Y \qquad \text{and}$$

$$\sigma_Y^2 > \sigma_X^2$$

This suggests that outcomes different from the mean are more likely for random variable **Y** than for random variable **X**.

comparing two different variables with the same mean --> use variance

❖ Useful Results for Expected Value and Variance

Let **a** and **b** be any constant fixed numbers.

- $E(a) = a$
- $E(a + bX) = a + b\,E(X) = a + b\,\mu_X$

  This result can be shown:

$$\begin{aligned} E(a + bX) &= \sum_x (a + bx)P(x) \\ &= a\sum_x P(x) + b\sum_x x\,P(x) \\ &= a + b\,E(X) \end{aligned}$$

- $Var(a) = 0$
- $Var(a + bX) = b^2\,Var(X)$

  This result can be shown:

$$\begin{aligned} Var(a + bX) &= E[\{a + bX - E(a + bX)\}^2] \\ &= E[\{a + bX - (a + b\mu_X)\}^2] \\ &= E[(bX - b\mu_X)^2] \\ &= E[b^2(X - \mu_X)^2] \\ &= b^2\,E[(X - \mu_X)^2] \\ &= b^2\,Var(X) \end{aligned}$$

a+bX linear function

a verschwindet weil
Varianz einer
Konstanten = 0

Example

A production process gives variation for the number of paper clips per package.

Let the random variable **X** be the number of paper clips in a package. The probability distribution function and cumulative probability function are given as:

x	P(x)	F(x)
47	0.04	0.04
48	0.13	0.17
49	0.21	0.38
50	0.29	0.67
51	0.20	0.87
52	0.10	0.97
53	0.03	1.00

Questions and answers.

- Two packages are chosen at random. Find the probability that at least one of them contains at least 50 paper clips.

  With the assumption of independence the answer is:

$$1 - P(\text{both contain less than 50}) = 1 - P(X \le 49)^2$$

$$= 1 - F(49)^2$$

$$= 1 - 0.38^2$$ square because two boxes

$$= 0.8556$$

- Find the mean and standard deviation of the number of paper clips per package.

The mean is calculated as:

$$E(X) = \sum_x x\,P(x)$$

$$= (47)(.04) + (48)(.13) + (49)(.21) + (50)(.29) + (51)(.20) + (52)(.10) + (53)(.03)$$

$$= 49.9$$

To calculate the variance, first calculate:

$$E(X^2) = \sum_x x^2\,P(x)$$

$$= 47^2(.04) + 48^2(.13) + 49^2(.21) + 50^2(.29) + 51^2(.20) + 52^2(.10) + 53^2(.03)$$

$$= 2491.96$$

The variance is found as:

$$\begin{aligned} Var(X) &= E(X^2) - \mu_X^2 \\ &= 2491.96 - 49.9^2 \\ &= 1.95 \end{aligned}$$

Varianz ist immer positiv!

The standard deviation is calculated as:

$$\sigma_X = \sqrt{1.95} = 1.396$$

- Now suppose that the cost (in cents) of producing a package of paper clips is the random variable:

$$\mathbf{C = 16 + 2X}$$

and the price of a package of paper clips is $1.50.

Therefore, profit (in cents) per package is the random variable:

$$\begin{aligned}\mathbf{P} &= \mathbf{150 - C} \\ &= \mathbf{134 - 2X}\end{aligned}$$

Find the mean and standard deviation of profit per package.

The mean is:

$$\begin{aligned}\mathbf{E(P)} &= \mathbf{E(134 - 2X)} \\ &= \mathbf{134 - 2E(X)} \\ &= \mathbf{134 - (2)(49.9) = 34.2} \quad \text{cents}\end{aligned}$$

The variance is:

$$\begin{aligned}\boldsymbol{\sigma_P^2} = \mathbf{Var(P)} &= \mathbf{Var(134 - 2X)} \\ &= \mathbf{(-2)^2\, Var(X)} \\ &= \mathbf{4}\boldsymbol{\sigma_X^2}\end{aligned}$$

This gives the standard deviation:

$$\boldsymbol{\sigma_P} = \sqrt{4\sigma_X^2} = 2\sigma_X = \mathbf{(2)(1.396) = 2.79} \quad \text{cents}$$

Chapter 4.4 Binomial Distribution

A special application of a discrete probability distribution is the binomial distribution.

To start, introduce the random variable X_B that takes two outcomes:

$x = 1$ "success"

$x = 0$ "failure"

The probability distribution function of X_B is:

$P(X_B = 1) = p$ for $0 < p < 1$ (the probability of success)

$P(X_B = 0) = 1 - p$

This is known as the Bernoulli distribution.
The mean and variance are calculated as:

$$\mu_{X_B} = E(X_B) = \sum_x x\,P(x)$$
$$= 1 \cdot p + 0 \cdot (1 - p)$$
$$= p$$

$$Var(X_B) = E(X_B^2) - \mu_{X_B}^2$$
$$= \sum_x x^2\,P(x) - p^2$$
$$= (1)(1) \cdot p + (0)(0) \cdot (1 - p) - p^2$$
$$= p - p^2$$
$$= p(1 - p)$$

Now consider that a random experiment with the outcome of success or failure is repeated **n** times.
Each trial produces success or failure with probabilities **p** and **(1–p)** respectively. Assume independence so that the result of one trial does not influence the result of any other trial.

Let the random variable **X** be the number of successes in **n** trials.

The probability distribution function of **X** is defined as:

$$\mathbf{P(x) = P(} \text{x successes in n independent trials} \mathbf{)}$$

for $\mathbf{x} = 0, 1, 2, \ldots, \mathbf{n}$

This is known as the **binomial distribution**.

To establish the properties of the binomial distribution, consider the random variables $X_1, X_2, \ldots, X_n$ as **n** independent Bernoulli variables each with probability of success **p**.

The total number of successes is: $X = \sum_{i=1}^{n} X_i$

The mean of the binomial random variable is then:

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = n \cdot p$$

and, with independence, the variance is:

$$Var(X) = \sum_{i=1}^{n} Var(X_i) = n \cdot p \cdot (1 - p)$$

A calculation formula for the binomial probabilities can be obtained as follows.

In **n** independent trials, the probability of **x** successes and (**n**–**x**) failures is:

$$p^x (1-p)^{n-x}$$

1-p = failure

The number of combinations of **x** successes in **n** trials is:

$$C_x^n = \frac{n!}{x!(n-x)!}$$

Therefore, the probability distribution function for the binomial distribution is:

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \qquad \text{for } x = 0, 1, 2, \ldots, n$$

Note: a calculation rule is $0! = 1$

4 trials:
P(X=1)=4*p(1-p)3

wie ist die Wahrscheinlichkeit, dass ich bei 4 versuchen einmal richtig liege

Examples of application of the binomial distribution are:

$$P(X=0) = P(\text{ no successes in n trials})$$

$$= (1-p)^n$$

$$P(X=1) = P(\text{ one success in n trials})$$

$$= n\,p\,(1-p)^{n-1}$$

The cumulative probability function of the binomial distribution may have useful application and is calculated as:

$$F(x) = P(X \le x) \qquad \text{for } x = 0, 1, 2, \ldots, n$$

For example,

$$F(1) = P(X \le 1) = P(\text{ at most one success in n trials})$$

$$= P(X=0) + P(X=1)$$

Example

A company installs new heating furnaces. For any installation, the probability of a return visit for a repair is 0.15.
Six installations are made in a given week.
Assume independence of outcomes for these installations.

Let the random variable **X** be the number of return visits.
X follows a binomial distribution with:

$$\mathbf{p} = 0.15 \quad \text{and} \quad \mathbf{1 - p} = 0.85$$

Find the probability that a return visit will be needed in more than one of the installations.

$$\begin{aligned} \mathbf{P(X > 1)} &= \mathbf{1 - P(\, X \le 1)} \\ &= \mathbf{1 - [P(X = 0) + P(X = 1)]} \\ &= \mathbf{1 - [}(0.85)^{\mathbf{6}} + (6)(0.15)(0.85)^{\mathbf{5}}\mathbf{]} \end{aligned}$$

At this point, it is clear that the numerical calculations can be tedious.

Numerical answers for binomial distribution probabilities can be obtained with Microsoft Excel.

Select Insert Function   BINOM.DIST

The general usage is: BINOM.DIST(x, n, p, cumulative)

where   cumulative = 0   for the probability distribution function,

cumulative = 1   for the cumulative probability function

For this exercise the calculation of $\mathbf{P(X \leq 1)}$ was found with

BINOM.DIST(1, 6, 0.15, 1) = 0.7765

Therefore, the answer is:

$$\mathbf{P(X > 1) = 1 - P(X \leq 1) = 1 - 0.7765 = 0.2235}$$

Note: an important assumption of the binomial distribution is independent trials.

Chapter 4.7 Jointly Distributed Random Variables

Economic relationships between variables are of interest.

Let $\mathbf{X}$ and $\mathbf{Y}$ be a pair of discrete random variables such that:

$\mathbf{X}$ has numerical outcomes x, and

$\mathbf{Y}$ has numerical outcomes y.

The **joint probability function** is:

$$\mathbf{P_{X,Y}(x,y) = P(X = x \text{ and } Y = y)} \qquad \text{for all pairs } (x, y)$$

A joint probability function has the properties:

- $\mathbf{0 \leq P_{X,Y}(x,y) \leq 1}$ for all pairs (x, y)
- $\mathbf{\sum_{x} \sum_{y} P_{X,Y}(x,y) = 1}$

The probability function of $\mathbf{X}$ is obtained by summing the joint probabilities:

$$\mathbf{P_X(x) = \sum_y P_{X,Y}(x,y)} \qquad \text{for all possible values of x.}$$

↑

summation over all possible values of y.

This is called the **marginal probability function** of $\mathbf{X}$.

Similarily, the marginal probability function of $\mathbf{Y}$ is constructed as:

$$\mathbf{P_Y(y) = \sum_x P_{X,Y}(x,y)} \qquad \text{for all possible values of y.}$$

The **conditional probability function** of $\mathbf{Y}$ given that $\mathbf{X} = x$ is:

$$\mathbf{P_{Y|X}(y \mid x) = \frac{P_{X,Y}(x,y)}{P_X(x)}} \qquad \text{for all possible values of y.}$$

Similarily, the conditional probability function of $\mathbf{X}$ given that $\mathbf{Y} = y$ is:

$$\mathbf{P_{X|Y}(x \mid y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}} \qquad \text{for all possible values of x.}$$

For the conditional probability function:

$$\mathbf{\sum_y P_{Y|X}(y \mid x) = 1} \qquad \text{and} \qquad \mathbf{\sum_x P_{X|Y}(x \mid y) = 1}$$

The random variables **X** and **Y** are **independent** if and only if:

$$P_{X,Y}(x,y) = P_X(x)\,P_Y(y) \qquad \text{for all pairs } (x, y)$$

If random variables **X** and **Y** are independent then:

$$P_{Y|X}(y \mid x) = \frac{P_{X,Y}(x,y)}{P_X(x)}$$

$$= \frac{P_X(x)\,P_Y(y)}{P_X(x)}$$

$$= P_Y(y)$$

That is, the conditional probability function of **Y**, given that the random variable **X** takes the value x, is identical to the marginal probability function of **Y**, for all possible values of y.

The mean and variance of a conditional distribution of a random variable can be obtained.

The conditional mean of **Y** given **X**=x is calculated as:

$$\mathbf{E(Y \mid X = x) = \sum_{y} y\, P_{Y|X}(y \mid x)}$$

The conditional variance of **Y** given **X**=x is:

$$\mathbf{Var(Y \mid X = x) \quad = E\left(\left(Y - E(Y \mid X = x)\right)^2 \mid X = x\right)}$$

$$\mathbf{= \sum_{y} (y - \mu_{Y|X})^2\, P_{Y|X}(y \mid x)}$$

where $\mathbf{\mu_{Y|X} = E(Y \mid X = x)}$

For calculation:

$$\mathbf{Var(Y \mid X = x) = E(Y^2 \mid X = x) - \mu_{Y|X}^2}$$

where $\mathbf{E(Y^2 \mid X = x) = \sum_{y} y^2\, P_{Y|X}(y \mid x)}$

Example

A survey by a real estate agent has collected information on apartment rentals. Consider the discrete random variables:

X volume of inquiries by renters. The possible values are:

x = 0 little interest

x = 1 moderate interest

x = 2 strong interest

Y number of lines in a newspaper ad. Possible values are:

y = 3, 4, 5

The joint probability function is:

	X		
Y	0	1	2
3	0.09	0.14	0.07
4	0.07	0.23	0.16
5	0.03	0.10	0.11

Questions and Answers

- Find the probability function of **X**.

$$\mathbf{P_X(0) =}\ 0.09 + 0.07 + 0.03 = 0.19$$

$$\mathbf{P_X(1) =}\ 0.14 + 0.23 + 0.10 = 0.47$$

$$\mathbf{P_X(2) =}\ 0.07 + 0.16 + 0.11 = 0.34$$

Note: the probabilities sum to one.

- Find the mean of **X**.

wie bei marginal probability function mit Gewichtung

$$\boldsymbol{\mu_X = E(X) = \sum_x x\, P_X(x)}$$

$$= (0)(0.19) + (1)(0.47) + (2)(0.34) = 1.15$$

- For the random variable **Y**, find the probability function and mean.

$$\mathbf{P_Y(3) =}\ 0.30,\quad \mathbf{P_Y(4) =}\ 0.46,\quad \text{and}\quad \mathbf{P_Y(5) =}\ 0.24$$

$$\boldsymbol{\mu_Y = E(Y) = \sum_y y\, P_Y(y)}$$

$$= (3)(0.30) + (4)(0.46) + (5)(0.24) = 3.94$$

- Find the conditional probability function for **Y** given **X**=0.

$$\mathbf{P_{Y|X}(3 \mid 0) = \frac{P_{X,Y}(0,3)}{P_X(0)}} = \frac{0.09}{0.19} = 0.4737$$

$$\mathbf{P_{Y|X}(4 \mid 0) = \frac{P_{X,Y}(0,4)}{P_X(0)}} = \frac{0.07}{0.19} = 0.3684$$

$$\mathbf{P_{Y|X}(5 \mid 0) = \frac{P_{X,Y}(0,5)}{P_X(0)}} = \frac{0.03}{0.19} = 0.1579$$

Note: the probabilities sum to one.

- Find the mean and variance of **Y** given **X**=0.

$\boldsymbol{\mu_{Y|X}}$ = (3)(0.4737) + (4)(0.3684) + (5)(0.1579) = 3.684

To calculate the conditional variance, first calculate:

$$\mathbf{E(Y^2 \mid X = 0) = \sum_y y^2 \, P_{Y|X}(y \mid x)}$$

$$= 3^2 (0.4737) + 4^2 (0.3684) + 5^2 (0.1579)$$

$$= 14.105$$

The conditional variance is:

$$\mathbf{Var(Y \mid X = 0) \quad = E(Y^2 \mid X = 0) - \mu_{Y|X}^2}$$

$$= 14.105 - 3.684^2 = 0.53$$

- Are **X** and **Y** independent ?

Recall that independence requires:

$$\mathbf{P_{X,Y}(x,y) = P_X(x)P_Y(y)} \qquad \text{for all pairs } (x, y)$$

For the values **X** = 0 and **Y** = 3, the joint probability is:

$$\mathbf{P_{X,Y}}(0,3) = 0.09$$

The product of the marginal probabilities is:

$$\mathbf{P_X}(0)\mathbf{P_Y}(3) = (0.19)(0.30) = 0.057$$

It is clear that $\mathbf{P_{X,Y}}(0,3) \neq \mathbf{P_X}(0)\mathbf{P_Y}(3)$

Therefore, the two random variables are not independent.

Let $\mathbf{g(X, Y)}$ be a function of the discrete random variables $\mathbf{X}$ and $\mathbf{Y}$. The **expected value** of this function is defined as:

$$E[g(X,Y)] = \sum_x \sum_y g(x,y) P_{X,Y}(x,y)$$

A property of expectation is:

$$E(X+Y) = E(X) + E(Y)$$

This result can be shown:

$$\begin{aligned}
E(X+Y) &= \sum_x \sum_y (x+y) P_{X,Y}(x,y) \\
&= \sum_x \sum_y \left[x P_{X,Y}(x,y) + y P_{X,Y}(x,y) \right] \\
&= \sum_x x \sum_y P_{X,Y}(x,y) + \sum_y y \sum_x P_{X,Y}(x,y) \\
&= \sum_x x P_X(x) + \sum_y y P_Y(y) \\
&= E(X) + E(Y)
\end{aligned}$$

For constant fixed numbers $\mathbf{a}$ and $\mathbf{b}$ a rule is:

$$E(aX + bY) = aE(X) + bE(Y)$$

A general result is that for $\mathbf{K}$ random variables $X_1, X_2, \ldots, X_K$ with means $\mu_1, \mu_2, \ldots, \mu_K$ the expected value of their sum is:

$$E(X_1 + X_2 + \ldots + X_K) = \mu_1 + \mu_2 + \ldots + \mu_K$$

A measure of a linear relationship between two random variables is of interest.

For random variables $\mathbf{X}$ and $\mathbf{Y}$ with means μ_X and μ_Y the **covariance** between $\mathbf{X}$ and $\mathbf{Y}$ is defined as:

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

uncorrelated = cov = 0

$$= \sum_x \sum_y (x - \mu_X)(y - \mu_Y) P_{X,Y}(x, y)$$

An equivalent expression can be stated:

1) calculate the mean: μl

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

$$= E[(X\,Y - X\,\mu_Y - \mu_X\,Y + \mu_X\,\mu_Y)]$$

$$= E(XY) - \mu_X\,\mu_Y - \mu_X\,\mu_Y + \mu_X\,\mu_Y$$

$$= E(XY) - \mu_X\,\mu_Y$$

where

$$E(X\,Y) = \sum_x \sum_y x\,y\,P_{X,Y}(x, y)$$

in general: E(x,y) =/μx*μy= E(x)*E(y)

when independent: E(x,y) = μx*μy

If the random variables X and Y are independent then:

$$\begin{aligned}
\mathbf{E(X\,Y)} &= \sum_x \sum_y x\,y\,P_{X,Y}(x,y) \\
&= \sum_x \sum_y x\,y\,P_X(x)\,P_Y(y) \\
&= [\sum_x x\,P_X(x)][\sum_y y\,P_Y(y)] \\
&= \mu_X\,\mu_Y
\end{aligned}$$

It follows that independence gives:

$$\mathbf{Cov(X,Y) = E(XY) - \mu_X\,\mu_Y = 0}$$

- However, if zero covariance is established, this does **not** guarantee that the random variables are independent. Covariance is designed to measure the possibility of a linear relationship. Nonlinear relationships between variables may give dependencies even though the covariance is zero.

Also note that, in general, for random variables with non-zero covariance:

$$\mathbf{E(XY) \neq E(X)\,E(Y)}$$

Covariance gives an indication of the sign (positive or negative) of a linear relationship between random variables.
A measure of the strength of a linear relationship between random variables $\mathbf{X}$ and $\mathbf{Y}$ is the **correlation** defined as:

$$\rho = \mathbf{Corr(X, Y)} = \frac{\mathbf{Cov(X, Y)}}{\sigma_{\mathbf{X}}\,\sigma_{\mathbf{Y}}}$$

↑
Greek letter rho

where $\sigma_{\mathbf{X}}$ and $\sigma_{\mathbf{Y}}$ are the standard deviations of the random variables.

A result is: $\mathbf{-1 \le \rho \le 1}$

A value of $\rho = 0$ indicates that the random variables are **uncorrelated**.

Problem: If two random variables are uncorrelated, are they independent ?

Example: the real estate agent exercise Continued.

Earlier in the lecture notes, an exercise introduced the joint probability function:

	X		
Y	0	1	2
3	0.09	0.14	0.07
4	0.07	0.23	0.16
5	0.03	0.10	0.11

To find the covariance between the random variables **X** and **Y** first calculate:

$$\mathbf{E(X\,Y) = \sum_{x}\sum_{y} x\,y\,P_{X,Y}(x,y)}$$

$$\begin{aligned} &= (0)(3)(0.09) + (0)(4)(0.07) + (0)(5)(0.03) + \\ &\quad (1)(3)(0.14) + (1)(4)(0.23) + (1)(5)(0.10) + \\ &\quad (2)(3)(0.07) + (2)(4)(0.16) + (2)(5)(0.11) \end{aligned}$$

$$= 4.64$$

The covariance is:

$$\mathbf{Cov(X,Y) = E(XY) - \mu_X\,\mu_Y}$$

$$= 4.64 - (1.15)(3.94)$$

$$= 0.109$$

not that important

The variances of the two random variables are calculated as:

$$\sigma_X^2 = E(X^2) - \mu_X^2$$

$$= (0)(0.19) + (1)(0.47) + (4)(0.34) - (1.15)(1.15)$$

$$= 0.5075$$

$$\sigma_Y^2 = E(Y^2) - \mu_Y^2$$

$$= (9)(0.30) + (16)(0.46) + (25)(0.24) - (3.94)(3.94)$$

$$= 0.5364$$

The correlation between **X** and **Y** is:

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{0.109}{\sqrt{(0.5075)(0.5364)}} = 0.2089$$

That is, there is a positive correlation between the number of lines in a newspaper ad and the volume of inquiries about the apartment rental.

❖ Useful Results for the Variance of a Linear Combination of Random Variables

For two random variables $\mathbf{X}$ and $\mathbf{Y}$ a result is:

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

This result can be shown:

$$\begin{aligned}
\mathbf{Var(X+Y)} \quad &= \mathbf{E[\{X + Y - E(X + Y)\}^2]} \\
&= \mathbf{E[\{(X - E(X)) + (Y - E(Y))\}^2]} \\
&= \mathbf{E[(X - E(X))^2 + (Y - E(Y))^2 + 2\,(X - E(X))(Y - E(Y))]} \\
&= \mathbf{E[(X - E(X))^2] + E[(Y - E(Y))^2] + 2\,E[(X - E(X))(Y - E(Y))]} \\
&= \mathbf{Var(X) \qquad\quad + Var(Y) \qquad\quad + 2\,Cov(X, Y)}
\end{aligned}$$

Another result is:

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

When $\mathbf{X}$ and $\mathbf{Y}$ are independent then the covariance is zero and:

$$\mathbf{Var(X + Y) = Var(X - Y) = Var(X) + Var(Y)}$$

For constant fixed numbers $\mathbf{a}$ and $\mathbf{b}$, a general rule is:

$$\mathbf{Var(aX + bY) = a^2\,Var(X) + b^2\,Var(Y) + 2\,a\,b\,Cov(X, Y)}$$

auch bei 0.5* Var (X+Y) einfach in die Klammer reinziehen
= Var (0.5*X + 0.5*Y)

Example: Portfolio Analysis

The job of a financial advisor may be to recommend a mix of stocks or a portfolio for investment purposes.

Consider the random variables:

X_1 price of one share of stock for company A

X_2 price of one share of stock for company B

X_3 price of one share of stock for company C

For Company A possible prices are \$45, \$50, \$55, \$60 and for Company B and C possible prices are \$40, \$50, \$60, \$70.
That is, the prices are styled to give discrete random variables.
A more realistic presentation is to view stock prices as a continuous random variable – the topic of the next chapter.

Using the joint probability function (not stated here) means and variances are:

$E(X_1) = \$\,53$ $\quad Var(X_1) = 31.3$

$E(X_2) = \$\,55$ $\quad Var(X_2) = 125$

$E(X_3) = \$\,55$ $\quad Var(X_3) = 125$

and covariances are:

$Cov(X_1, X_2) = 59.17$ (positive covariance) and

$Cov(X_1, X_3) = -\,59.17$ (negative covariance)

Two alternative portfolios are represented by the random variables:

$$W_1 = 5X_1 + 10X_2 \quad \text{and}$$

$$W_2 = 5X_1 + 10X_3$$

The mean of the first portfolio is calculated as:

$$\begin{aligned} E(W_1) &= E(5X_1 + 10X_2) \\ &= 5E(X_1) + 10E(X_2) \\ &= (5)(53) + (10)(55) \\ &= \$815 \end{aligned}$$

The portfolio W_2 has the identical mean as the portfolio W_1.
That is,

$$E(W_2) = E(W_1)$$

The performance of the two portfolios can be compared by comparing their variances.

$$\mathbf{Var(W_1)} = \mathbf{Var(5X_1 + 10X_2)}$$

$$= \mathbf{5^2\, Var(X_1) + 10^2\, Var(X_2) + 2(5)(10) \cdot Cov(X_1, X_2)}$$

$$= (25)(31.3) + (100)(125) + (100)(59.17)$$

$$= 19{,}199.5$$

In contrast:

$$\mathbf{Var(W_2)} = \mathbf{5^2\, Var(X_1) + 10^2\, Var(X_3) + 2(5)(10) \cdot Cov(X_1, X_3)}$$

$$= (25)(31.3) + (100)(125) + (100) \cdot (-59.17)$$

$$= 7{,}365.5$$

High variance implies high risk.

negative cov lowers risk

The effect of the negative covariance is to reduce the variance and hence to reduce the risk of the portfolio.