L5

Basic Statistics for Economists

Spring 2020

Department of Statistics

Stockholm University

Last time

Combinatorics

- Choose with or without replacement
- Does order matter or does order not matter? $(0! = 1)$
- Permutations and combinations (faculty e.g. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$)

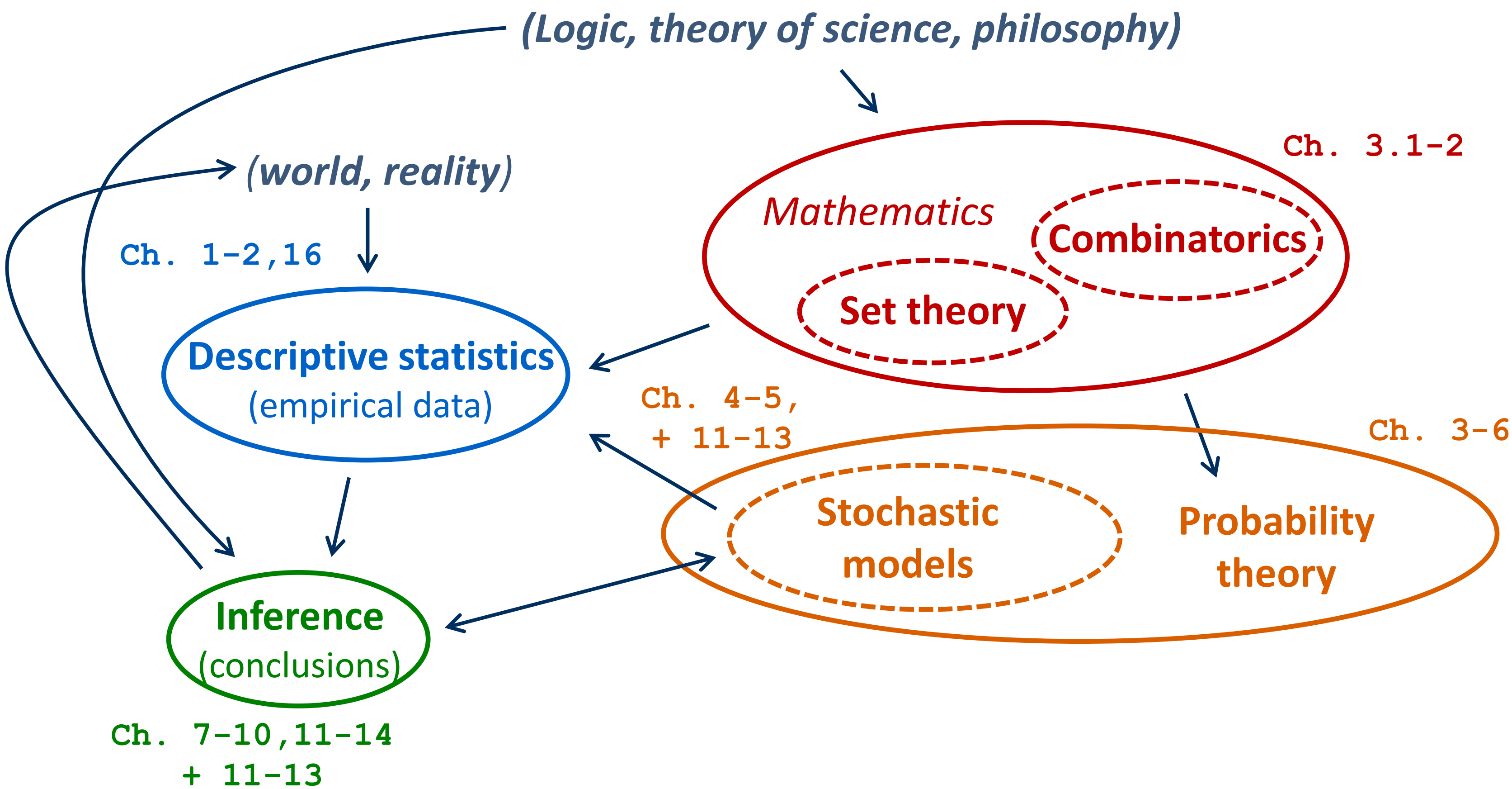
	Order matters	Order does not matter
With replacement	$\boldsymbol{n^x}$	$\left(C_x^{n+x-1} \right)$
Without replacement	$\boldsymbol{P_x^n = \frac{n!}{(n-x)!}}$	$\boldsymbol{C_x^n = \binom{n}{x} = \frac{n!}{x!\,(n-x)!}}$

Last time

Probability Theory

- Interpretations of probability
- The postulates in NCT (Kolmogorov's axioms)
- Calculation rules
  - Complement rule $P(\bar{A}) = 1 - P(A)$
  - Addition rule $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - Conditional probability $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
  - Multiplication rule $P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$
  - Independence $P(A \cap B) = P(A) \cdot P(B)$

Previously



Stockholm University

Today

- ***Probability Theory***
  - Bayes' theorem
- ***Random variables (models)***
  - What is a random variable? (abbreviated r.v.)
  - Probability distributions for discrete r.v.
  - Expectation and variance of r.v.
  - Linear transformations
  - A particularly common distribution:

    The Binomial Distribution

Stockholm University

More on conditioning

- The probability of an intersection maybe unknown, but possible to calculate:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Multiplication rule
L4 p. 27

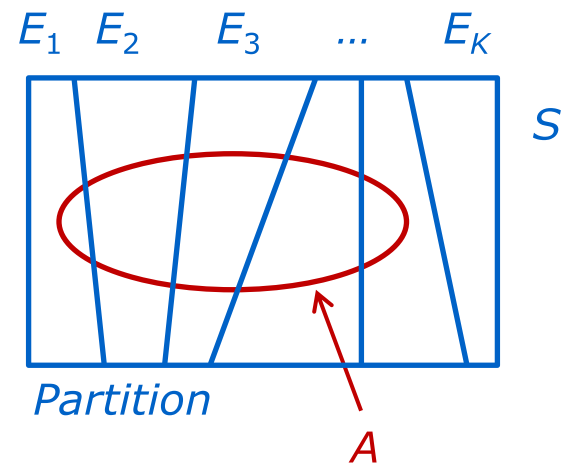
- Hence, one can calculate the ***reversed*** *conditional probabilities*:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B) \cdot P(B)}{P(A)}$$

- It is common that stated probabilities are conditional, even though it might not be obvious

Stockholm University

The Theorem of Total Probability



- If the events $E_1, E_2, \dots, E_K$ are *disjoint*
- and $E_1, E_2, \dots, E_K$ *cover* S:

$$E_1 \cup E_2 \cup \cdots \cup E_K = S$$

- Then the collection $E_1, E_2, \dots, E_K$ is called a ***partition*** of S.
- Consider an event A and a partition $E_1, E_2, \dots, E_K$

$$\mathbf{P}(\boldsymbol{A}) = \mathrm{P}(A \cap E_1) + \mathrm{P}(A \cap E_2) + \cdots + \mathrm{P}(A \cap E_K)$$

$$= \mathrm{P}(A|E_1)P(E_1) + \mathrm{P}(A|E_2)P(E_2) + \cdots + \mathrm{P}(A|E_K)P(E_K)$$

$$= \sum_i \mathbf{P}(\boldsymbol{A}|\boldsymbol{E_i})\mathbf{P}(\boldsymbol{E_i})$$

Stockholm University

Exercise: burglary in single family homes

- An insurance company, *Småhusförsäkring,* insures homes in Solna, Danderyd och Ekerö. **What is the probability that a random customer gets their home burglarized.**
- The probability that an insured home is burglarized during 2019 is
  - Solna: 5%
  - Danderyd: 3%
  - Ekerö: 0,5%
- Where the customers live (no customer owns more than one home):
  - 10% live in Solna (E_1)
  - 30% live in Danderyd (E_2)
  - 60% live in Ekerö (E_3)

$\{E_1, E_2, E_3\}$ is a partition of the sample space of all the company's customers. Do you see why?

Stockholm University

Solution

- Let A be the event that a given customer's home is burglarized during 2019. The probability that a randomly chosen customer is subject to burglary is then:

$$P(A) = \sum_{i=1}^{3} P(A \mid E_i)\, P(E_i)$$

$$= P(A \mid E_1)\, P(E_1) + P(A \mid E_2)\, P(E_2) + P(A \mid E_3)\, P(E_3)$$
$$= 0{,}05 \cdot 0{,}10 + 0{,}03 \cdot 0{,}30 + 0{,}005 \cdot 0{,}60 = 0{,}017 = 1{,}7\%$$

Stockholm University

Bayes' theorem (3.15) p. 135

- The events $E_1 \cdots E_K$ are disjoint and cover all of S
- You also need an event A
- You have all the probabilities $P(E_i), \ldots, P(E_k)$, and the conditional probabilities $P(A|E_1), \ldots, P(A|E_k)$
- **Bayes theorem**:

$$P(E_i|A) = \frac{P(E_i \cap A)}{P(A)} = \frac{P(A|E_i)P(E_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{\sum_k P(A|E_k)P(E_k)}$$

Stockholm University

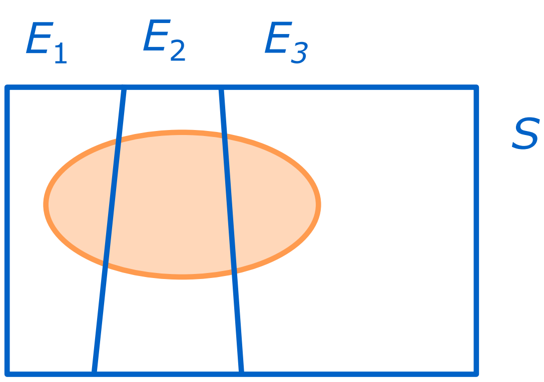
Exercise, Bayes' theorem

- Returning to the burglary example:
- The probability of burglary during 2018 is
  - Solna: 5%
  - Danderyd: 3%
  - Ekerö: 0,5%
- The customers:
  - 10% live in Solna (E_1)
  - 30% live in Danderyd (E_2)
  - 60% live in Ekerö (E_3)
- One of the customers reports the first burglary of the year. **What is the probability that they live in Solna?**

Solution

$$P(E_i|A) = \frac{P(E_i \cap A)}{P(A)} = \frac{P(A|E_i)P(E_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{\sum_k P(A|E_k)P(E_k)}$$

- Earlier, we found that $P(A) = 0{,}017$
- $P(A \mid E_1) = P(burglary| \, the \ customer \ lives \ in \ Solna) = 0{,}05$
- $P(E_1) = P(customer \ lives \ in \ Solna) = 0{,}1$
- $P(E_1|A) = \frac{P(A|E_3) \, P(E_3)}{P(A)} = \frac{0{,}05 \cdot 0{,}1}{0{,}017} \approx 0{,}29$



Not to scale

Random Variables

Also: ***stochastic variables***

- Suppose we have an experiment, the possible outcomes of which are ***numbers***
- Variables = something that varies
- Random variables = something that varies randomly
- Denoted by capital letters:

$$\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \dots$$

Stockholm University

Notation

Earlier:

- Capital letters $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, … denoted events
- Probabilities: $\boldsymbol{P(A)}$, $\boldsymbol{P(A \cup B)}$, $\boldsymbol{P(A|B)}$, …

With the notation for a r.v. $\boldsymbol{X}$ we may describe the events in a compact way:

Conditional Probability

- Events: "$\boldsymbol{X = 2}$", "$\boldsymbol{X > 3}$", "$\boldsymbol{X \neq 4}$" , "$\boldsymbol{X < 2 \cup X > 5}$" …
- **Probability of events:**

$$\boldsymbol{P(X=2)},\ \ \boldsymbol{P(X<2 \cup X>3)},\ \ \boldsymbol{P(X<2|X<3)}\ \ldots$$

- Instead of $\boldsymbol{P}(\mathit{The\ die\ shows\ less\ than\ two\ or\ more\ than\ three})$ …

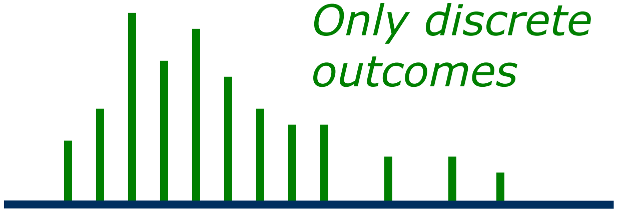
Stockholm University

Random variables, continued

- Two types

Random variables

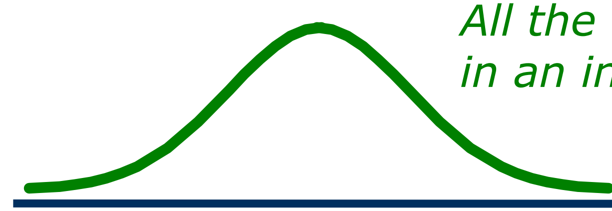
Discrete



Only discrete outcomes

The bars represent the probabilities of the different outcomes (L5, today)

Continuous



All the outcomes in an interval

The curve represents … ?

(L6, next time)

Stockholm University

Example of a discrete random variable

A customer buys $\boldsymbol{n} = \mathbf{10}$ units of some product

Let $\boldsymbol{X}$ = ”the number out of the $\mathbf{10}$ that malfunctions within a year”

- Sample space: $\boldsymbol{S_X} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots, \mathbf{10}\}$
- $\boldsymbol{S_X}$ = elementary disjoint outcomes of the experiment (numbers)
- We may define events as we wish:

$A = \{0, 1, 2\}$	$0 \leq X \leq 2$	$X < 3$
$\bar{A} = \{3, \dots, 10\}$	$2 < X \leq 10$	$X \geq 3$
$B = \{0, 1, \dots, 9\}$	$\bar{B} = \{10\}$	$X = 10$
$C = \{11\}$	$11 \notin S$	

Notice! For discrete r.v. it matters whether the inequality is strict or not.

Stockholm University

The probability distribution function

- Or probability function or probability distribution or just the distribution *(abbreviated: **pdf**)*
- Denoted $\boldsymbol{P(X=x)}$ or $\boldsymbol{P(x)}$
- If we have many r.v., we may clarify which one you we referring to by using indices, e.g. $P_X(x)$, $P_Y(y)$

$\boldsymbol{P_X(x)}$ represents the probability that X takes the value x. $\boldsymbol{P_X(x)}$ is a function of x.

We will write: $\boldsymbol{P(X=5)} = \boldsymbol{P_X(5)}$. Lower case x is replaced by a number. The function evaluated at that number gives us a probability.

Stockholm University

The probability function, continued

Conditions: the probability of all disjoint, elementary outcomes $\boldsymbol{x} \in \boldsymbol{S_X}$, given by $\boldsymbol{P_X}(\boldsymbol{x})$ has to be ≥ 0 and sum up to 1:

$$P_X(x) \geq 0 \; \forall x \in S_X \quad \text{and} \quad \sum_{x \in S} P_X(x) = 1$$

cumulative distribution function (abbreviated cdf):

$$\boldsymbol{F_X(x_0) = P(X \leq x_0)} = \sum_{x \leq x_0} P_X(x)$$

Compare to the cumulative frequencies from F2.

The probability that $\boldsymbol{X}$ does not exceed the value $\boldsymbol{x_0}$. $\boldsymbol{F_X}$ is a function of $\boldsymbol{x_0}$.

Stockholm University

Example

- Consider the following probability function:

x	1	2	3	4	Summa
$P_X(x)$	0,25	0,40	0,25	0,10	1,00
$F_X(x)$	0,25	0,65	0,90	1,00	

Calculate:

a) $P(X = 3) = [\text{read directly from the table}] = 0{,}25$

b) $P(X \leq 3) = F_X(3) = [\text{from the table}] = 0{,}90$

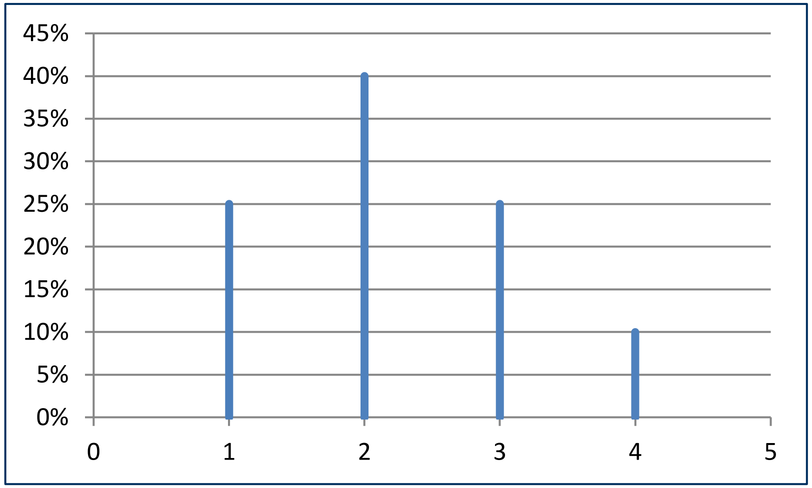
c) $P(X \geq 3) = P_X(3) + P_X(4) = 0{,}25 + 0{,}10 = 0{,}35\ = 1 - F_X(2) = 1 - 0{,}65$

d) $P(X > 3) = P_X(4) = 0{,}10$

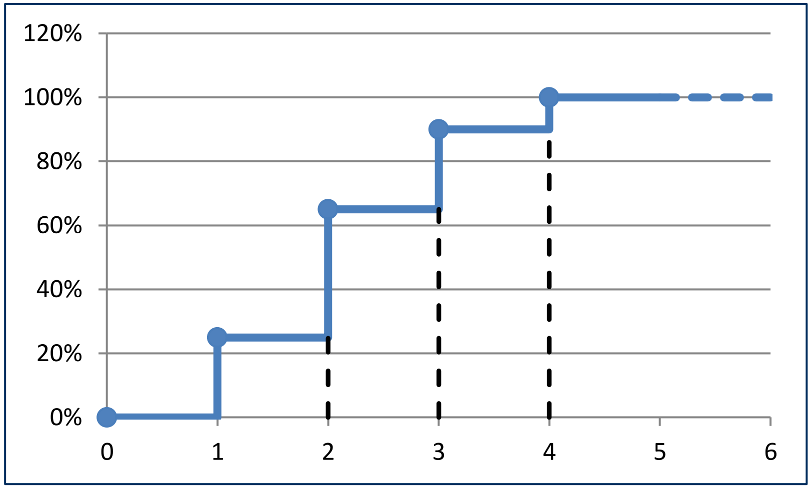
Graphically

- A random variable $\boldsymbol{X}$ with $\boldsymbol{S} = \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}\}$:

The probability function
$\boldsymbol{P_X(x) = P(X = x)}$



The cumulative probability function
$\boldsymbol{F_X(x_0) = P(X \leq x_0)}$



Stockholm University

Summary measures

- The probability function $P_X(x)$ contains all necessary information about X.
  - *Assumed but not stated: we know what S_X is.*
- **Summary measures** describe some important attributes of a probability function.
- ***"Location" – expected value***
  - theoretical average, what we can expect the average to be
- ***"Spread" – variance***
  - How spread out the values are

The expected value of a discrete r.v.

Expected value is defined as: $\mu_X = E[X] = \sum_{x \in S_X} x P_X(x)$

- Expected value, mean, average
- The expected value is a ***weighted mean*** of the elements in $\boldsymbol{S_X}$ where the weights are the probability of each element.
- Compare to the empirical mean in F2:

$$\bar{x} = \sum_{k=0}^{C} \left(k \cdot \frac{n_k}{n} \right)$$

Proportion with value k

What happens as $n \to \infty$?

Stockholm University

Example: The expectation of a Minitriss scratch card

The contribution to the expected value of the 5000 kr winning tickets

Expected value	Number of tickets	Win	P(Win = x)	x·P(Win=x)
	1	100 000	0,000001	0,1
	60	5 000	0,00006	0,3
	3000	100	0,003	0,3
	30000	30	0,03	0,9
	110000	20	0,11	2,2
	110000	10	0,11	1,1
	746939	0	0,746939	0
sum	**1000000**		**1**	**4,9**

The probability of winning 5000kr is calculated as the number of tickets with the prize 5000kr (60), divided by the total number of tickets (one million)

A not-yet-scratched minitriss card has an expected value of 4,90 kr. The cost of a minitriss is 10 kr, so your expected net win if you buy a ticket is: 4,90 – 10 = -5,10.

Stockholm University

Variance of a discrete r.v.

Variance is defined as: $$\sigma_X^2 = Var(X) = \sum_{x \in S_X} (x - \mu_X)^2 P_X(x)$$

- A weighted average of the squared distance to the expected value
- Compare to the empirical variance: $$s_x^2 = \sum_{k=0}^{C} (k - \bar{x})^2 \frac{n_k}{n-1}$$
- Alternative formula: $\sigma_X^2 = E(X^2) - \mu_X^2 = \left(\sum_{x \in S_X} x^2 P_X(x) \right) - \mu_X^2$

Stockholm University

Example

- Return of investment Y in thousands of dollars

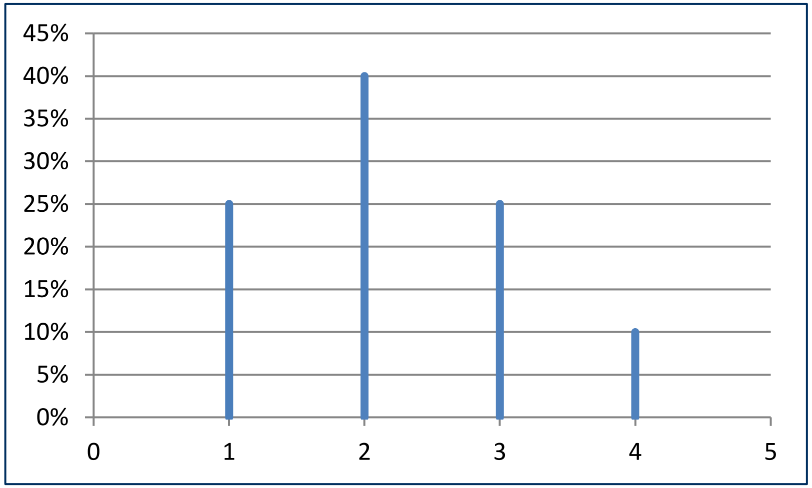
y	$P(y)$	$yP(y)$	$y-\mu$	$(y-\mu)^2$	$(y-\mu)^2P(y)$
-3	0,25	-0,75	-5	25	6,25
2	0,50	1,00	0	0	0
7	0,25	1,75	5	25	6,25
Sum	**1,00**	**2**	0	-	**12,5**

- $\mu_Y = E(Y) = -3 \cdot 0{,}25 + 2 \cdot 0{,}50 + 7 \cdot 0{,}25 = 2$
- $\sigma_Y^2 = Var(Y) = (-3-2)^2 \cdot 0{,}25 + (2-2)^2 \cdot 0{,}50 + (7-2)^2 \cdot 0{,}25 = 12{,}5$

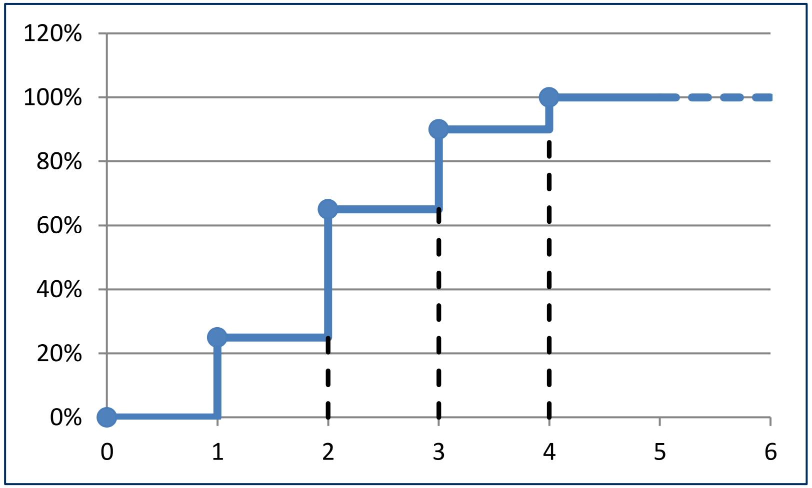
Previous example

- A random variable $\boldsymbol{X}$ with $\boldsymbol{S} = \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}\}$:

Probability function
$\boldsymbol{P_X(x) = P(X = x)}$



Cumulative probability function
$\boldsymbol{F_X(x_0) = P(X \leq x_0)}$



Stockholm University

Previous example

- The same pdf as in the previous example:

x	1	2	3	4	Sum
$P(x)$	0,25	0,40	0,25	0,10	1,00
$xP(x)$	0,25	0,80	0,75	0,40	**2,20**

$\mu_X = 2{,}20$

$x-\mu$	$-1{,}20$	$-0{,}20$	0,80	1,80	
$(x-\mu)^2$	1,44	0,04	0,64	3,24	
$(x-\mu)^2P(x)$	0,36	0,016	0,16	0,324	**0,86**

$\sigma_X^2 = 0{,}86$

x^2	1	4	9	16	
$x^2P(x)$	0,25	1,60	2,25	1,60	5,70

$$\sigma_X^2 = E(X^2) - \mu_X^2 = 5{,}70 - 2{,}20^2 = 0{,}86$$

The expected value of a function of a r.v.

- We can calculate the expected value of any function $g(x)$:

  $$E[g(X)] = \sum_x g(x) P(x)$$

  $g(x)$ instead of just x

- Example:

  $$E[\ln(X)] = \sum_x \big(\ln(x) \cdot P(x)\big)$$

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

- But observe that in general $\boldsymbol{E[g(X)] \neq g(E[X]) = g(\mu)}$

Stockholm University

Special case – Linear combinations

- A function of the type $\boldsymbol{Y = g(X) = a + bX}$ is called a **linear combination**

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

  Then:

$$\mu_Y = E(Y) = E(a + bX) = \boldsymbol{a + b\mu_X}$$

$$\sigma_Y^2 = Var(Y) = Var(a + bX) = \boldsymbol{b^2\sigma_X^2}$$

Notice! $\boldsymbol{b}$ squared, $\boldsymbol{b^2}$

- If you have a linear function $Y = g(X)$ you may calculate expected value and variance of Y directly.

Stockholm University

Example

X = temperature in Celsius and Y = temperature in Fahrenheit

- Linear function: $Y = 32 + \frac{9}{5}X$ $a = 32$ and $b = \frac{9}{5}$

- Suppose that the average temperature is $\mu_X = 20$ and $\sigma_X^2 = 5$

  $\mu_Y = a + b\mu_X = 32 + \frac{9}{5} \cdot 20 = \mathbf{68}$ $\sigma_Y^2 = \boldsymbol{b^2}\sigma_X^2 = \frac{81}{25} \cdot 5 = \mathbf{16,2}$

Stockholm University

Standardization

Standardization **IMPORTANT!**

Suppose
$$\boldsymbol{Z} = \frac{\boldsymbol{X} - \boldsymbol{\mu_X}}{\boldsymbol{\sigma_X}} = -\frac{\mu_X}{\sqrt{\sigma_X^2}} + \frac{1}{\sqrt{\sigma_X^2}} \cdot X$$

$$\boldsymbol{\mu_Z} = -\frac{\mu_X}{\sqrt{\sigma_X^2}} + \frac{1}{\sqrt{\sigma_X^2}} \cdot \mu_X = \mathbf{0} \qquad \boldsymbol{\sigma_Z^2} = \boldsymbol{b^2} \sigma_X^2 = \left(\frac{1}{\sqrt{\sigma_X^2}}\right)^2 \cdot \sigma_X^2 = \mathbf{1}$$

Bernoulli experiment

- A simple experiment with only ***two*** possible outcomes:

  Yes or No, Success or Failure, …

- Let $X = 1$ if success and $X = 0$ if failure; $S_X = \{0, 1\}$
- Suppose that $P(X = 1) = P_X(1) = P$ where $0 \leq P \leq 1$

$\boldsymbol{x}$	$\boldsymbol{P_X(x)}$	$\boldsymbol{xP_X(x)}$	$\boldsymbol{x - \mu}$	$\boldsymbol{(x - \mu)^2}$	$\boldsymbol{(x - \mu)^2 P_X(x)}$
1	P	P	$1 - P$	$(1 - P)^2$	$(1 - P)^2 P$
0	$1 - P$	0	$-P$	P^2	$P^2(1 - P)$

- $\boldsymbol{\mu_X} = 1 \cdot P + 0 \cdot (1 - P) = \boldsymbol{P}$
- $\boldsymbol{\sigma_X^2} = (1 - P)^2 P + P^2(1 - P) = P(1 - P)(1 - P + P) = \boldsymbol{P(1 - P)}$

Stockholm University

Bernoulli distribution

- $\boldsymbol{P}$ is called the **parameter** of the distribution
  - The value of P determines everything: probabilities, expected value and variance
  - We get a family of distributions depending on the choice of $\boldsymbol{P}$

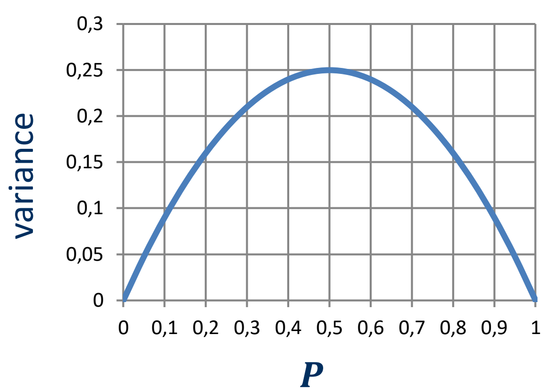
- For different values of $\boldsymbol{P}$, what is $\boldsymbol{\sigma_X^2 = P(1-P)}$:

$$\boldsymbol{P = 0 \Rightarrow P(1-P) = 0 \quad P = 1 \Rightarrow P(1-P) = 0}$$

- For what value of $\boldsymbol{P}$ is $\boldsymbol{\sigma_X^2}$ greatest?
  - Set the derivative w.r.t. $\boldsymbol{P}$ to zero

$$\frac{\partial}{\partial P}P(1-P) = 1 - 2P = 0 \Rightarrow \boldsymbol{p = 0.5}$$

$$\Rightarrow P(1-P) = \boldsymbol{0,25}$$



Stockholm University

The probability function for Bernoulli

- A summary formula, a function of x
- Calculate the probabilities for all the possible outcomes of the Bernoulli distribution
- $\boldsymbol{P(x) = P^{x}(1 - P)^{1-x}}$ for $x = 0 \text{ eller } 1$
- Plug in: $x = 0 \Rightarrow \quad P^{0}(1 - P)^{1-0} = 1 \cdot (1 - P) = \mathbf{1} - \boldsymbol{P}$

  $x = 1 \Rightarrow \quad P^{1}(1 - P)^{1-1} = P \cdot 1 = \boldsymbol{P}$

- ***This confirms what we knew. Not exciting, but wait! This will get more exciting…***

Stockholm University

Many Bernoulli experiments

- Imagine that we repeat the Bernoulli experiment $\boldsymbol{n}$ times
- **Independent** trials and the **same value of $\mathbf{P}$** each time

IMPORTANT! IMPORTANT!

- For each trial, we record whether the result was 0 or 1
- $X_i = 1$ if success and $X_i = 0$ if failure for trial number i, $i = 1, \dots n$
- Define r.v. $X =$ ”number of 1’s after n trials”
- Observe that $\boldsymbol{X} = \boldsymbol{X_1} + \boldsymbol{X_2} + \cdots + \boldsymbol{X_n}$
- What is the distribution of X? How do you calculate $P(X = x)$?

Stockholm University

Binomial distribution

Let X = the sum of n independent Bernoulli distributed r.v. with constant value for P.

Then X is **Binomially distributed** r.v. with pdf:

$$\boldsymbol{P(x) = \binom{n}{x} P^x (1-P)^{n-x}}$$

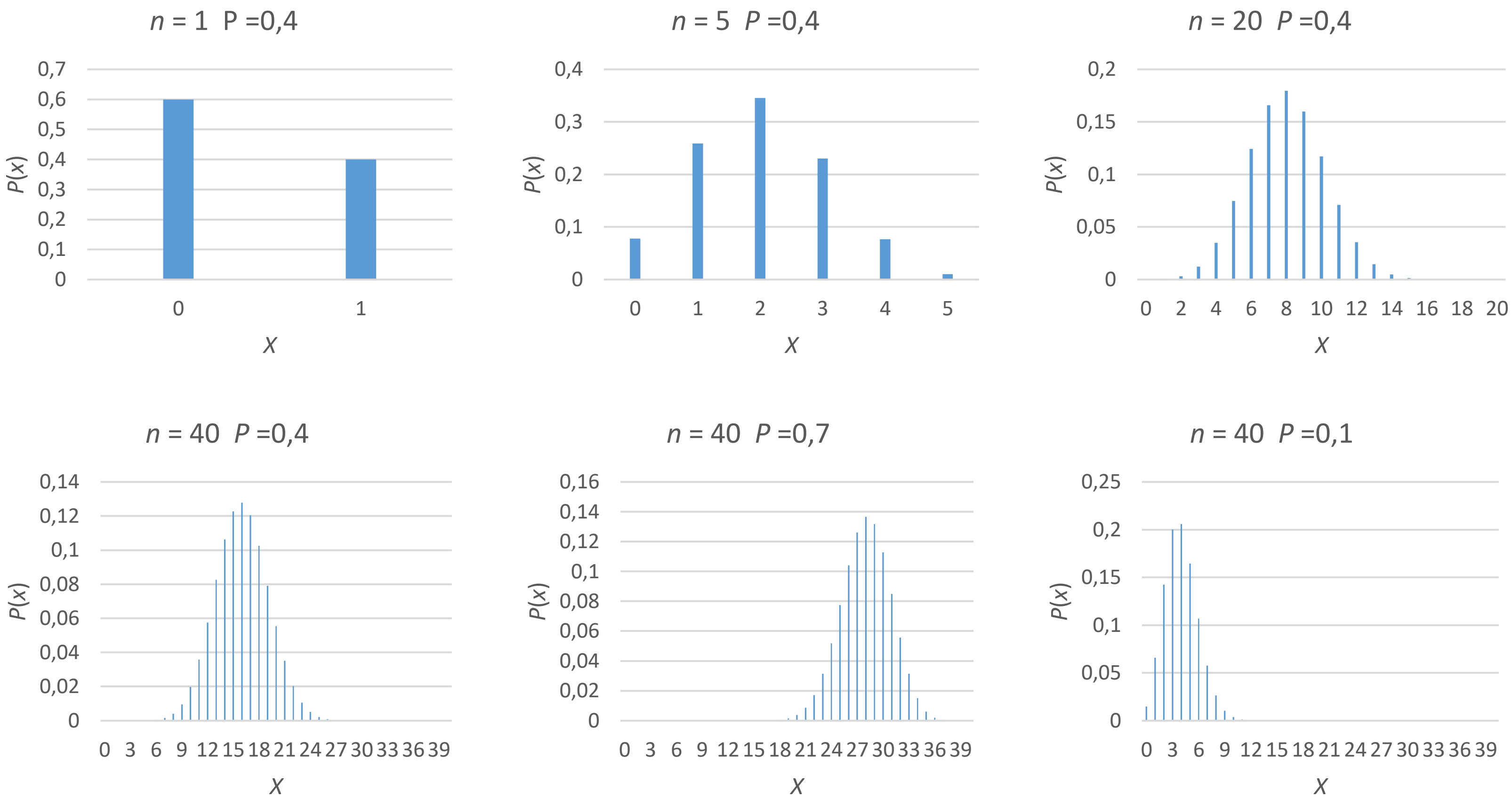
The probability $\boldsymbol{P}$ and the number of trials $\boldsymbol{n}$ is the **parameters** of the distribution. We write $\boldsymbol{X \sim Bin(n, P)}$.

Sample space: $\boldsymbol{S_X = \{0, 1, 2, \dots, n\}}$

Expected value: $\boldsymbol{\mu_X = nP}$ Variance: $\boldsymbol{\sigma_X^2 = nP(1-P)}$

Stockholm University

Some Binomial distributions



Example

In a population of elderly people, 30% suffer from high blood pressure. We randomly choose $n = 5$ people from this population and we let X be the number of people from our sample who were found to have high blood pressure.

- The variable X is binomially distributed with $n = 5$ and $P = 0{,}3$.
- X can take the values $0,1,2,3,4$ or 5.
- What is the probability that $X = 2$, i.e. $P(X = 2)$?
  - i.e. what is the probability of two 1's and three 0's?
- What is the probability of $X \leq 2$, i.e. $P(X \leq 2)$?

Example: using the probability function

- Random variable: $X \sim Bin(5; 0,3)$
- Parameter values: $n = 5;\ p = 0,3$
- Sample space: $S_X = \{0,1,2,3,4,5\}$
- Probability function: $P(x) = \binom{5}{x} \cdot 0,3^x \cdot 0,7^{5-x}$
- Plugging in values, we get:

$\mu = np\ = 1,5$
$\sigma^2 = np(1-p) = 1,05$

1. $\boldsymbol{P(2)} = \binom{5}{2} 0,3^2 0,7^3 = 10 \cdot 0,09 \cdot 0,343 = \mathbf{0,30870}$

2. $\boldsymbol{P(X \leq 2)} = F(2) = P(0) + P(1) + P(2)$

$= \binom{5}{0} 0,3^0 0,7^5 + \binom{5}{1} 0,3^1 0,7^4 + \binom{5}{2} 0,3^2 0,7^3$

$= 1 \cdot 1 \cdot 0,16807 + 5 \cdot 0,3 \cdot 0,2401 + 10 \cdot 0,09 \cdot 0,343 = \mathbf{0,83692}$

Another way: use the table

- Random variable $X \sim Bin(5; 0{,}3)$
- Parameter values: $n = 5 \;\; p = 0{,}3$

- Identify where in the table you can find the values

Notice that the table gives us values for $\boldsymbol{P(X \leq x)} = F(x)$

not for $P(X = x) = P(x)$

For a value of $P(x)$, find $F(x)$ and $F(x-1)$ and use

$$P(x) = F(x) - F(x-1)$$

Stockholm University

TABELL 7. Binomial-fördelningen; $n = 2 - 9$

$P(X \leq x)$ där $X \in Bin(n, p)$. För $p > 0,5$, utnyttja att $P(X \leq x) = P(Y \geq n\text{-}x)$ där $Y \in Bin(n, 1\text{-}p)$

n	x	p = 0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
2	0	0,90250	0,81000	0,72250	0,64000	0,56250	0,49000	0,42250	0,36000	0,30250	0,25000
	1	0,99750	0,99000	0,97750	0,96000	0,93750	0,91000	0,87750	0,84000	0,79750	0,75000
3	0	0,85738	0,72900	0,61413	0,51200	0,42188	0,34300	0,27463	0,21600	0,16638	0,12500
	1	0,99275	0,97200	0,93925	0,89600	0,84375	0,78400	0,71825	0,64800	0,57475	0,50000
	2	0,99988	0,99900	0,99663	0,99200	0,98438	0,97300	0,95713	0,93600	0,90888	0,87500
4	0	0,81451	0,65610	0,52201	0,40960	0,31641	0,24010	0,17851	0,12960	0,09151	0,06250
	1	0,98598	0,94770	0,89048	0,81920	0,73828	0,65170	0,56298	0,47520	0,39098	0,31250
	2	0,99952	0,99630	0,98802	0,97280	0,94922	0,91630	0,87352	0,82080	0,75852	0,68750
	3	0,99999	0,99990	0,99949	0,99840	0,99609	0,99190	0,98499	0,97440	0,95899	0,93750
5	0	0,77378	0,59049	0,44371	0,32768	0,23730	0,16807	0,11603	0,07776	0,05033	0,03125
	1	0,97741	0,91854	0,83521	0,73728	0,63281	0,52822	0,42842	0,33696	0,25622	0,18750
	2	0,99884	0,99144	0,97339	0,94208	0,89648	0,83692	0,76483	0,68256	0,59313	0,50000
	3	0,99997	0,99954	0,99777	0,99328	0,98438	0,96922	0,94598	0,91296	0,86878	0,81250
	4	1,00000	0,99999	0,99992	0,99968	0,99902	0,99757	0,99475	0,98976	0,98155	0,96875
6	0	0,73509	0,53144	0,37715	0,26214	0,17798	0,11765	0,07542	0,04666	0,02768	0,01563
	1	0,96723	0,88574	0,77648	0,65536	0,53394	0,42018	0,31908	0,23328	0,16357	0,10938
	2	0,99777	0,98415	0,95266	0,90112	0,83057	0,74431	0,64709	0,54432	0,44152	0,34375
	3	0,99991	0,99873	0,99411	0,98304	0,96240	0,92953	0,88258	0,82080	0,74474	0,65625

Example: using the table

2. Read from the table:

$$P(X \leq 2) = F(2) = \mathbf{0,83692}$$

1. Read from the table:

$$P(X = 2) = P(X \leq 2) - P(X \leq 1) = [Draw!]$$

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

$$= 0{,}83692 - 0{,}52822 = \mathbf{0,30870}$$

Summary

- Bayes' theorem – to flip the conditioning
- Random variable, discrete random variables
- Expected value and variance
  - Theoretical versions of the mean and variance of a sample
- Linear combinations and standardizing - important!
- Bernoulli experiments – the simple case
- Binomial distribution – very common and useful
  - When we can use it
  - its parameters, expected value and variance
  - How you calculate probabilities

Some things that we did not mention

The table can be used for $\boldsymbol{P \leq 0,50}$.

- What if $\boldsymbol{P > 0,5}$?
- Read section 4.4, we will return to this during the exercises.

  Tip:
  - Define $\boldsymbol{Y = n - X}$ = number of failures, number of 0's.
  - What is the distribution of Y?

The table can be used for $\boldsymbol{n \leq 20}$.

- What if $\boldsymbol{n > 20}$?
- Read section 5.4, we will return to this during L8

Stockholm University

Exercise (do at home)

- In a corporation, 60% of the employees are positive to a particular change.
- Suppose that we have a random sample of four people and that we wish to measure the number of people who are positive to the change in our sample.

a) Construct the probability function for r.v. $X =$ ”Number of positive people in the sample.”

b) Calculate expected value and variance of X.

c) What is the probability of (exactly) two positive in the sample?

d) What is the probability of at least two positive?

Next time

Continuous random variables

- The probability function and the density function
- Expected values, variances, linear combinations

The Normal distribution

- The most common and most useful continuous distribution
- How to use the normal table and calculate probabilities

- Seeing Theory: https://seeing-theory.brown.edu/