

UPPSALA UNIVERSITET

FÖRELÄSSNINGSANTECKNINGAR

Sannolikhetssteori 2

Rami Abou Zahra

Inlämningsdatum
September 1, 2023

CONTENTS

1. Repetition	2
1.1. Probability measure	2
1.2. Conditional probability theory	3
1.3. Independent events	3
1.4. Random variables	4
2. Transformations	6

1. REPETITION

Anmärkning: Det rekommenderas starkt att läsa igenom anteckningarna från Sannolikhetsteori 1

Definition/Sats 1.1: Random trial

An event is not certain, it usually has a probability associated with it. Taking that "risk" to see what the outcome is, is called a random trial.

Examples of random trials include throwing dice, picking cards, number of people who pass a road

Different possibilities (outcomes).

In the example of the dice, the outcomes are 1-6

Definition/Sats 1.2: Events

An event is something that happens (or does not happen) when you the random trial

You can have an event based on one outcome, or multiple.

Example (one outcome): The dice is 3 after a throw

Example (several outcomes): The card is 7 or lower (1,2,3,4,5,6, all the different colours)

Example (0 outcomes): The card shows both spades and hearts at the same time (impossible)

1.1. Probability measure.

Related to the probability that an event occurs.

Definition/Sats 1.3: Probability measure

A *probability measure* is a function which satisfies Kolmogorovs axioms and for each event gives a number $\in [0, 1]$

The number is called the *probability* of the event. Usually denoted $P = P(A)$ where $A \subset \Omega$

Definition/Sats 1.4: Kolmogorovs axioms

Let $P : 2^\Omega \rightarrow \mathbb{R}$. P is called a *probability measure* if it satisfies the following

- $P(A) \geq 0 \quad \forall A \in 2^\Omega$
- $P(\Omega) = 1$
- $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \quad (A_i \text{ disjoint})$

1.2. Conditional probability theory.

One can think of conditional probability in the term of Venn Diagrams in order to create intuition. Usually, if one event has happened it will affect other events and it is of interest to take this into consideration when calculating the probability of events.

The probability of an event A occurring given that the event B has occurred is denoted by

$$P(A|B)$$

Example:

Let A be the event that a person has 2 daughters, let B be the event that a person has 0 daughters, and C be the event that he has at least 1 daughter.

The probability $P(A|B)$ is of course 0, since given that he has 0 daughters, the probability is 0 for him to have 2 at the same time as he has 0

$P(B|C)$ is also 0, using similar argument as above

$P(A|C) = \frac{P(A \cap C)}{P(C)}$ We cannot say much here, other than that the probability is strictly positive since we already have one child

Definition/Sats 1.5: Bayes theorem

Let F_1, \dots, F_n be disjoint events $\in \Omega$ with $P(F_i) > 0$, and $P(\bigcup F_i) = 1$.

$$P(F_j|E) = \frac{P(E|F_j) \cdot P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}$$

Example:

Suppose we have 3 different cards. The first card is red on both sides (RR), the second card is black on both sides (BB), and the third card is black and red (RB)

We draw a card at random of these three cards such that we only see one side of the card. Now suppose the side we see is red, what is the probability that the other side is black?

We are interested in the event $P(RB|R)$:

$$\begin{aligned} \frac{P(RB \cap R)}{P(R)} &= \underbrace{\quad}_{\text{Bayes}} = \frac{P(R|RB)P(RB)}{P(R|RR)P(RR) + P(R|RB)P(RB) + P(R|BB)P(BB)} \\ &= \frac{(1/2)(1/3)}{1 \cdot (1/3) + (1/2) \cdot (1/3) + (0) \cdot (1/3)} = \frac{1}{3} \end{aligned}$$

1.3. Independent events.

Definition/Sats 1.6: Independent events

If $P(A|B) = P(A)$, then A and B are independent

Example:

Let A be the event that 2 parents get a daughter, and B be the event that the neighbors child ate an ice cream yesterday.

Since these events do not affect each other, they are independent.

Example:

Let A be the event that the first throw of a dice yields 6, and let B be the event that the second throw is 3. Then A and B are independent since the first throw does not affect the second throw:

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

Anmärkning:

There is an equivalent definition for independence through the following:

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

Anmärkning:

Independence is a symmetric relationship.

1.4. Random variables.

Definition/Sats 1.7: Random variable

A *random variable* is a function that for each outcome associates a number with it.

An example is a person's age, or the value of a card drawn. If the outcome is random, the number is also random.

Each random variable has a distribution function associated with it, and is defined as $F(X) = P(X \leq x)$

Anmärkning:

$$\begin{aligned} \lim_{X \rightarrow -\infty} F(X) &= 0 \\ \lim_{X \rightarrow \infty} F(X) &= 1 \end{aligned}$$

If $X_1 < X_2 \Rightarrow F(X_1) \leq F(X_2)$

We also have that F is right-continuous, meaning

$$\lim_{X \rightarrow a^+} F(X) = F(a)$$

There are 2 types of random variables that will be covered in this course, discrete and continuous (there are also absolutely continuous random variables, but they will not be covered)

1.4.1. Discrete random variables.

Definition/Sats 1.8: Discrete random variables

Consists of a finite or countable infinite set of numbers with probabilities:

- $P(X = x_i) = P(x_i) > 0$
- $P(X = \bigcup_{i=1}^{\infty} x_i) = 1$

Anmärkning:

If we have an uncountable infinite set of possibilities, the probability would be 0. Here is where continuous variables come to play

1.4.2. Continuous random variables.

For a continuous random variable, $F(x)$ is differentiable so that there exists a function f such that:

$$F(x) = \int_{-\infty}^x f(t)dt$$

From this comes 2 important things we can derive (both from the discrete and continuous case), namely expected value and the variance

Definition/Sats 1.9: Expected value

For discrete random variables, it is defined as

$$E(X) = \sum xF(x_i)$$

For the continuous case:

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

Definition/Sats 1.10: Variance

$Var(X) = E(X - E(X))^2$ for both discrete and continuous random variables

An equivalent definition is $E(X^2) - (E(X))^2$

Anmärkning:

$E(X^2)$ is called the second moment

Anmärkning:

If the variance is small, we know that the random variable does not fluctuate a lot from the expected value.

If X is a random variable (r.v) with density function f and g is a function, then we can define a new random variable $g(X) = Y$

Y is a random variable with density function \hat{f} .

Then:

$$E(Y) = \int y\hat{f}(y)dy = E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

And for the discrete random variable we have:

$$E(g(X)) = \sum g(x_i)p(x_i)$$

Anmärkning:

It is often better to use the definition of the density function for X rather than Y

Another remark worth noting is that $E(X^2)$ is a special case of $\int g(x)f(x)dx$

Example:

This example considers a distribution with no expected value (∞), and therefore it has no variance.

$P(X = k) = \frac{1}{k} - \frac{1}{k+1}$, this fulfills Kolmogorovs axioms, and

$$E(X) = \sum_{k=1}^{\infty} \frac{k}{k} - \frac{k}{k+1} = \sum_{k=1}^{\infty} \left(1 - \frac{k}{k+1}\right) = \sum_{k=1}^{\infty} \frac{1}{k+1} = \infty$$

2. TRANSFORMATIONS

2.1. Pre-knowledge.

Let X_i be independent random variables with the same mean value (expected value) μ and variance σ^2

Let $S_n = \sum_{i=1}^n X_i$ $\xRightarrow{\text{Law of large numbers}} \frac{S_n}{n} \rightarrow \mu$ (convergence in probability).

Of course, this is assuming some sort of equal distribution.

The notation for convergence in probability is denoted by $Y_n \xrightarrow{P} a$. This follows from Markov's inequality. It is strongly suggested to look in the notes from the first course here.

From the law of large numbers, $S_n \approx n\mu$. But this does not take into account some errors that may take place in the $\frac{S_n}{n}$ side, as this does not affect the convergence to μ .

This is treated with the *Central Limit Theorem* (CLT), which says that $S_n \sim N(n\mu, n\sigma^2)$

This is equivalent to say that:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \approx N(0, 1)$$

When n is large.

Anmärkning:

Here we talked about *convergence in distribution*