Probability and Inferential Statistics

Vorlesung SS 16

Distribution

Prof. Dr. Rolf Steyer

Distribution

The definition of a random variable X on a probability space $(\Omega, \mathscr{A}, P)$ with values in $(\Omega', \mathscr{A}')$ implies that all inverse images

$$X^{-1}(A') := \{\omega \in \Omega \colon X(\omega) \in A'\}, \quad A' \in \mathscr{A}',$$

are elements of the σ-algebra $\mathscr{A}$ on Ω. Because the measure $P : \mathscr{A} \to [0, 1]$ assigns a probability to *all* elements of $\mathscr{A}$, the probabilities $P[X^{-1}(A')]$ of these inverse images are determined by P.

Definition 1. Let $X \colon (\Omega, \mathscr{A}, P) \to (\Omega', \mathscr{A}')$ be a random variable. Then the function $P_X \colon \mathscr{A}' \to [0,1]$ defined by

$$P_X(A') \;=\; P[X^{-1}(A')], \quad \forall\, A' \in \mathscr{A}', \tag{1}$$

is called the *distribution* of X (with respect to P).

Example: Indicator Variable

Let $(\Omega, \mathscr{A}, P)$ be a probability space and $A \in \mathscr{A}$. Then $\mathit{1}_A$ is called the *indicator (variable) of (the event) A*, if

$$\mathit{1}_A(\omega) \;:=\; \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The inverse image of the sets $\{1\}$ and $\{0\}$ under $\mathit{1}_A$ are

$$\mathit{1}_A^{\;-1}(\{1\}) \;=\; \{\mathit{1}_A = 1\} \;=\; A \qquad \text{and} \qquad \mathit{1}_A^{\;-1}(\{0\}) \;=\; \{\mathit{1}_A = 0\} \;=\; A^c.$$

Hence, the values of the distribution P_{1_A} are

$$P_{\mathit{1}_A}(\{1\}) = P(A), \quad P_{\mathit{1}_A}(\{0\}) = P(A^c), \quad P_{\mathit{1}_A}(\Omega') = P(\Omega) = 1, \quad P_{\mathit{1}_A}(\emptyset) = P(\emptyset) = 0.$$

Notation

Notation

If $A' \in \mathscr{A}'$, we use the notation

$$P(X \in A') \;:=\; P[X^{-1}(A')] \tag{3}$$

for the probability of the event $\{X \in A'\} = X^{-1}(A')$, i. e., the event that X takes on a value in the subset A' of Ω'. If $\{x\} \in \mathscr{A}'$, then we use

$$P(X{=}x) \;:=\; P[X^{-1}(\{x\})] \tag{4}$$

for the probability of the event $\{X{=}x\} = X^{-1}(\{x\}) = \{\omega \in \Omega\colon X(\omega) = x\}$. If we write $P(X{=}x)$, then we always assume $\{x\} \in \mathscr{A}'$, even if not mentioned explicitly.

New Probability Space

New Probability Space

The definition of a random variable implies that *every* random variable $X\colon (\Omega, \mathscr{A}, P) \to (\Omega', \mathscr{A}')$ has a distribution $P_X\colon \mathscr{A}' \to [0,1]$.

Furthermore, P_X is also a probability measure and $(\Omega', \mathscr{A}', P_X)$ is also a probability space.

Example

Consider again tossing a coin two times and the random variable X “number of heads”. The values

$$\begin{aligned} P_X(\{0\}) &= P[X^{-1}(\{0\})] = P[\{(t,t)\}] = 1/4, \\ P_X(\{1\}) &= P[X^{-1}(\{1\})] = P[\{(h,t),(t,h)\}] = 1/2, \\ P_X(\{2\}) &= P[X^{-1}(\{2\})] = P[\{(h,h)\}] = 1/4 \end{aligned}$$

of the distribution P_X are the probabilities of the events that X takes on the values 0, 1, and 2, respectively. Furthermore, the values

$$\begin{aligned} P_X(\{0,1\}) &= P[X^{-1}(\{0,1\})] = P(\{(t,t),(h,t),(t,h)\}] = 3/4, \\ P_X(\{0,2\}) &= P[X^{-1}(\{0,2\})] = P(\{(t,t),(h,h)\}) = 1/2, \\ P_X(\{1,2\}) &= P[X^{-1}(\{1,2\})] = P(\{(h,t),(t,h),(h,h)\}] = 3/4 \end{aligned}$$

of P_X are the probabilities of the events that X takes on the values in the set $\{0,1\}, \{0,2\}$, and $\{1,2\}$, respectively. The value $P_X(\Omega') = P[X^{-1}(\Omega')] = P(\Omega) = 1$ of the distribution P_X is the probability of the sure event and $P_X(\emptyset) = P[X^{-1}(\emptyset)] = P(\emptyset) = 0$ is the value of P_X for the impossible event.

Distribution Function

If we consider a univariate real-valued random variable X, then the *distribution function* F_X assigns to each $x \in \mathbb{R}$ the probability $P(X \le x)$ of the event $\{X \le x\} = \{\omega \in \Omega\colon X(\omega) \le x\}$ that X takes on a value *smaller or equal* than x. The distribution function uniquely determines the distribution P_X.

Definition 2. Let $X\colon (\Omega, \mathscr{A}, P) \rightarrow (\mathbb{R}, \mathscr{B}, P_X)$ denote a real-valued random variable with distribution P_X. Then the *(cumulative) distribution function* $F_X\colon \mathbb{R} \rightarrow [0,1]$ of X is defined by:

$$F_X(x) := P_X(]-\infty, x]) = P(X \le x), \quad \forall\, x \in \mathbb{R}. \tag{5}$$

F_X assigns to each $x \in \mathbb{R}$ the probability to take on a value *less than or equal to* x.

Probabilities and Intervals

The definition of a distribution function implies that we can compute the probability $P(a{<}X{\leq}b)$ of X taking a value in the interval $]a,b]$ by

$$P(a{<}X{\leq}b) \;=\; F_X(b) - F_X(a), \quad \text{if } a<b, \tag{6}$$

because

$$P(a{<}X{\leq}b) \;=\; P_X\big(\,]-\infty,b] \setminus\;]-\infty,a]\big) \;=\; P_X(]-\infty,b]) \;-\; P_X(]-\infty,a]).$$

Discrete Random Variable and Probability Function

The distribution of a discrete random variable can be described by its *probability function* that is now introduced. Remember, if $X\colon (\Omega,\mathscr{A},P) \to (\Omega',\mathscr{A}')$ is a random variable, then the distribution P_X of X is a probability measure on $(\Omega',\mathscr{A}')$.

Definition 3 (Discrete Random Variable and its Probability Function)**.** Let $X\colon (\Omega,\mathscr{A},P) \to (\Omega',\mathscr{A}')$ be a random variable and assume that $\Omega'_0 \subset \Omega'$ is finite or countable with $P_X(\Omega'_0) = 1$ and $\{x\} \in \mathscr{A}'$ for all $x \in \Omega'_0$. Then X and its distribution P_X are called *discrete*, and the function $p_X : \Omega' \to [0,1]$ defined by

$$p_X(x) \;=\; \begin{cases} P_X(\{x\}), & \text{if } x \in \Omega'_0, \\ 0, & \text{if } x \in \Omega' \setminus \Omega'_0, \end{cases} \tag{7}$$

is called the *probability function of X*.
We also use the notation $P(X{=}x) = p_X(x)$.

The distribution P_X assigns its values to each element of $\mathscr{A}'$, which are subsets of Ω'. In contrast, the probability function p_X assigns its values to each $x \in \Omega'_0$.

Example 1 - Flipping 2 Coins

Example

Consider again tossing a coin two times and the random variable X “number of heads”. Then

$$\begin{aligned} p_X(0) &= P_X(\{0\}) = 1/4, \\ p_X(1) &= P_X(\{1\}) = 1/2, \\ p_X(2) &= P_X(\{2\}) = 1/4 \end{aligned}$$

are the values of p_X. They are the probabilities of the events that X takes on the value 0, 1, and 2, respectively. These probabilities are also denoted $P(X{=}0)$, $P(X{=}1)$, $P(X{=}2)$, respectively.

Example 2 - Binomial Distribution

Let $(\Omega, \mathscr{A}, P)$ be a probability space and $X := \sum_{i=1}^{n} 1_{A_i}$, where 1_{A_i} is the indicator variable of $A_i \in \mathscr{A}$, $i = 1, \dots, n$. Assume

(a) $P(A_i) = p$ for all $i = 1, \dots, n$, and

(b) the events $A_1, \dots, A_n$ are independent.

Then the probability that X takes on the value x is

$$P(X{=}x) = b_{n,p}(x) = \binom{n}{x} p^x (1-p)^{n-x}, \tag{8}$$

where $x = 0, 1, \dots, n$ and

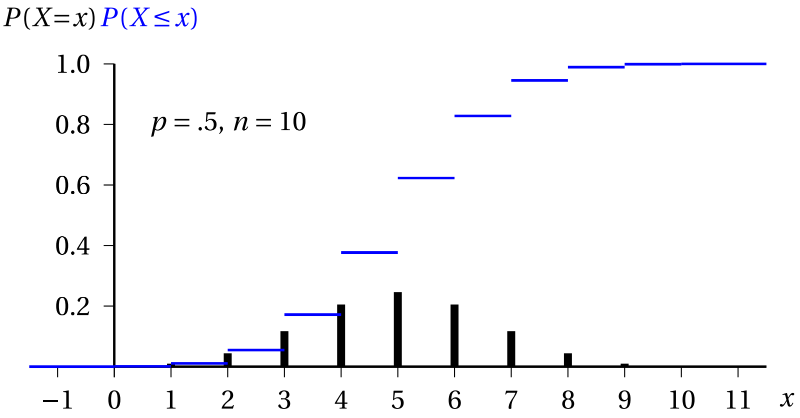
$$\binom{n}{x} := \frac{n!}{x! \cdot (n-x)!} \tag{9}$$

is the *binomial coefficient.* Note that $0! := 1$ and $x! := x \cdot (x-1) \cdot \ldots \cdot 1$ denotes the *factorial of* x.

If (8) holds, then P_X is called the *binomial distribution.*

An application is flipping a coin n times, where X is the random variable “number of heads”.

Graph of the Binomial Distribution



Probability function and distribution function of a binomial distribution

Continuous Random Variable and its Density

Definition 4. Let $X : \Omega \to \mathbb{R}$ be a real-valued random variable on $(\Omega, \mathscr{A}, P)$, and F_X its distribution function. Then X is called *continuous*, if there is a Riemann-integrable function $f_X : \mathbb{R} \to \mathbb{R}$, such that

$$F_X(x) := \int_{-\infty}^{x} f_X(t)\, dt, \quad x \in \mathbb{R}.$$

A function f_X satisfying this equation is called a *density* of X.

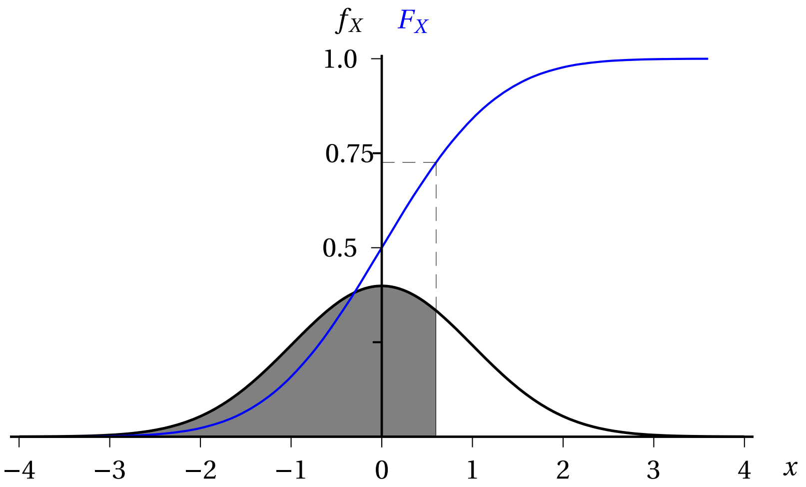
Example 3: Normal Distribution

Beispiel 1. A well-known density of a random variable X is the density of a *normal distribution*, which is defined by

$$f_X(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \cdot exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right),$$

where μ and σ^2 are real numbers. (Later on we will see that μ is the *expectation* of X and σ^2 its *variance*.)

Density and Distribution Function of the Normal Distribution



Density and distribution function of a normal distribution