Bernoulli distribution

From Wikipedia, the free encyclopedia

In probability theory and statistics, the **Bernoulli distribution**, named after Swiss scientist Jacob Bernoulli,[1] is the probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability $q = 1 - p$—i.e., the probability distribution of any single experiment that asks a yes–no question; the question results in a boolean-valued outcome, a single bit of information whose value is success/yes/true/one with probability p and failure/no/false/zero with probability q. It can be used to represent a coin toss where 1 and 0 would represent "head" and "tail" (or vice versa), respectively. In particular, unfair coins would have $p \neq 0.5$.

The Bernoulli distribution is a special case of the binomial distribution where a single experiment/trial is conducted (n=1). It is also a special case of the **two-point distribution**, for which the outcome need not be a bit, i.e., the two possible outcomes need not be 0 and 1.

Bernoulli

Parameters	$0 < p < 1, p \in \mathbb{R}$
Support	$k \in \{0, 1\}$
pmf	$\begin{cases} q = (1-p) & \text{for } k = 0 \\ p & \text{for } k = 1 \end{cases}$
CDF	$\begin{cases} 0 & \text{for } k < 0 \\ 1 - p & \text{for } 0 \le k < 1 \\ 1 & \text{for } k \ge 1 \end{cases}$
Mean	p
Median	$\begin{cases} 0 & \text{if } q > p \\ 0.5 & \text{if } q = p \\ 1 & \text{if } q < p \end{cases}$
Mode	$\begin{cases} 0 & \text{if } q > p \\ 0, 1 & \text{if } q = p \\ 1 & \text{if } q < p \end{cases}$
Variance	$p(1-p) (= pq)$
Skewness	$\frac{1-2p}{\sqrt{pq}}$
Ex. kurtosis	$\frac{1-6pq}{pq}$
Entropy	$-q\ln(q) - p\ln(p)$
MGF	$q + pe^t$
CF	$q + pe^{it}$
PGF	$q + pz$
Fisher information	$\frac{1}{p(1-p)}$

Contents

Properties of the Bernoulli Distribution

If X is a random variable with this distribution, we have:

$$\Pr(X = 0) = 1 - \Pr(X = 1) = 1 - p = q.$$

The probability mass function f of this distribution, over possible outcomes k, is

$$f(k; p) = \begin{cases} p & \text{if } k = 1, \\ 1 - p & \text{if } k = 0. \end{cases}$$

This can also be expressed as

$$f(k; p) = p^k (1-p)^{1-k} \quad \text{for } k \in \{0, 1\}.$$

The Bernoulli distribution is a special case of the binomial distribution with $n = 1$.[2]

The kurtosis goes to infinity for high and low values of p, but for $p = 1/2$ the two-point distributions including the Bernoulli distribution have a lower excess kurtosis than any other probability distribution, namely −2.

The Bernoulli distributions for $0 \le p \le 1$ form an exponential family.

The maximum likelihood estimator of p based on a random sample is the sample mean.

Mean

The expected value of a Bernoulli random variable X is

$$\mathrm{E}(X) = p$$

This is due to the fact that for a Bernoulli distributed random variable X with $\Pr(X = 1) = p$ and $\Pr(X = 0) = q$ we find

$$\mathrm{E}[X] = \Pr(X = 1) \cdot 1 + \Pr(X = 0) \cdot 0 = p \cdot 1 + q \cdot 0 = p$$

Variance

The variance of a Bernoulli distributed X is

$$\mathrm{Var}[X] = pq = p(1-p)$$

We first find

$$\mathrm{E}[X^2] = \Pr(X = 1) \cdot 1^2 + \Pr(X = 0) \cdot 0^2 = p \cdot 1^2 + q \cdot 0^2 = p$$

From this follows

$$\mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = p - p^2 = p(1-p) = pq$$

Skewness