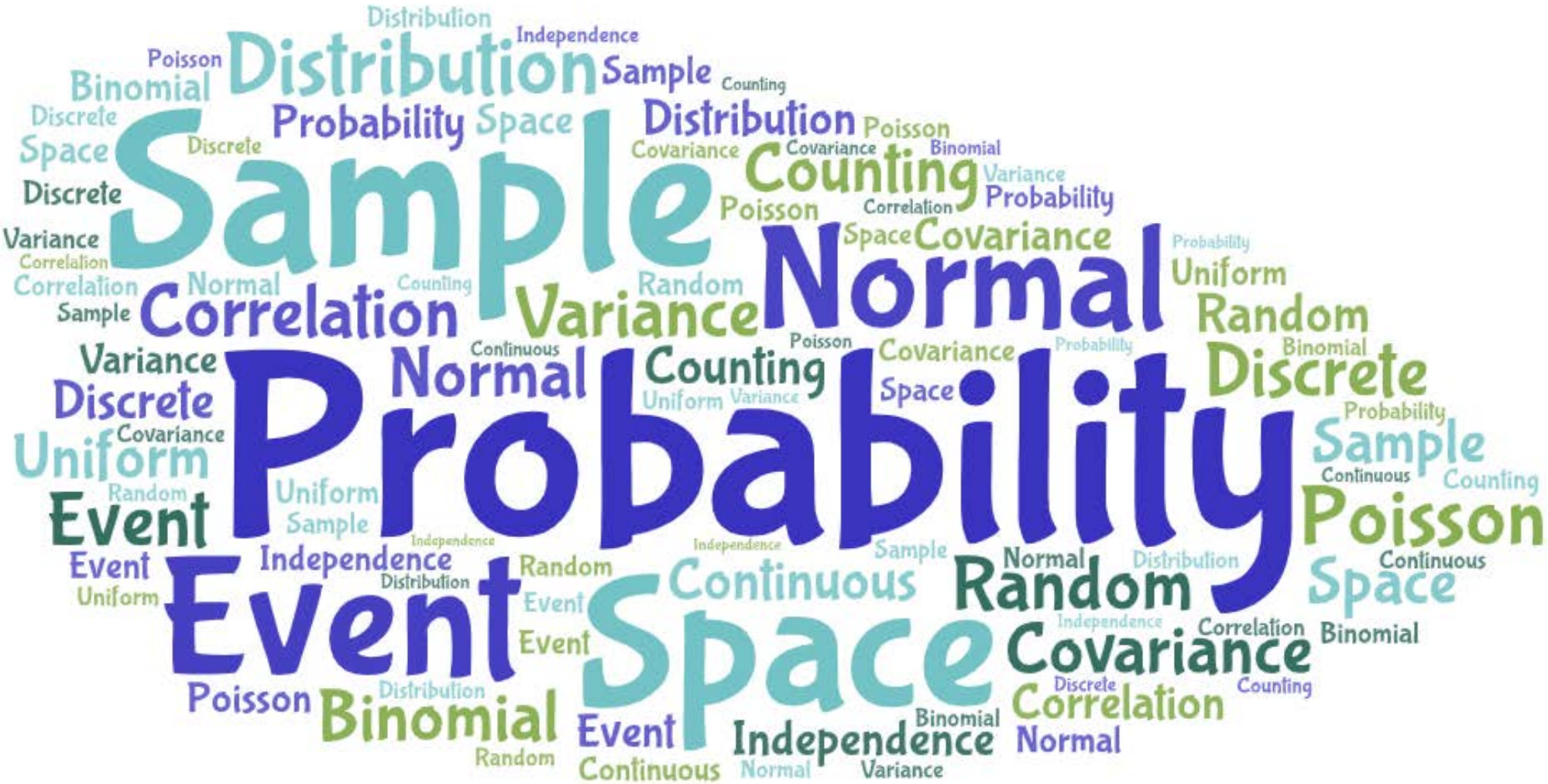
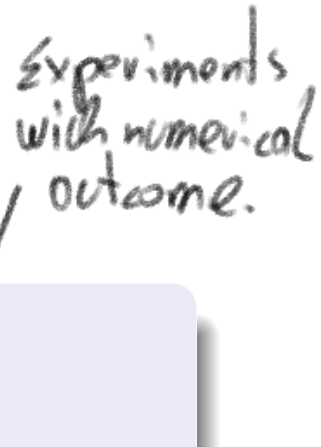
Discrete Mathematics and Probability

Week 8



Markulf Chris Heunen Kohlweiss

Topics



- Independence: what information changes probability
- Random variables: when variables depend on chance
- Expectation: most likely outcomes of experiment
- Variance: how much the experiment can deviate

Independence

Independence

Sometimes partial information on an experiment
does not change the likelihood of an event.

Definition

Events E and F are *independent* if $\mathbf{P}(E \mid F) = \mathbf{P}(E)$.
Equivalently: $\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F)$.
Equivalently: $\mathbf{P}(F \mid E) = \mathbf{P}(F)$.

Proposition

If E and F are independent events,
then E and F^c are also independent.

Independence

Sometimes partial information on an experiment does not change the likelihood of an event.

Definition

Events E and F are *independent* if $\mathbf{P}(E \mid F) = \mathbf{P}(E)$.
Equivalently: $\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F)$.
Equivalently: $\mathbf{P}(F \mid E) = \mathbf{P}(F)$.

Proposition

If E and F are independent events, then E and F^c are also independent.

Always give you some example propositions.

independent events $\neq$ mutually exclusive events

independence usually either trivial or tricky

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

E' = {sum is 7}

F = {first is 3}

$\frac{1}{36} = P(E' \cap F) = P(E') \cdot P(F) = \frac{6}{36} \cdot \frac{1}{6}$

independent!

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

E' = {sum is 7}

F = {first is 3}

G = {second is 4}

$\frac{1}{36} = P(E' \cap F) = P(E') \cdot P(F) = \frac{6}{36} \cdot \frac{1}{6}$

independent!

pairwise independent

$\frac{1}{36} = P(E' \cap G) = P(E') \cdot P(G) = \frac{6}{36} \frac{1}{6}$

$\frac{1}{36} = P(F \cap G) = P(F) \cdot P(G) = \frac{1}{6} \cdot \frac{1}{6}$

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

E' = {sum is 7}

F = {first is 3}

G = {second is 4}

$\frac{1}{36} = P(E' \cap F) = P(E') \cdot P(F) = \frac{6}{36} \cdot \frac{1}{6}$

independent!

pairwise independent

$\frac{1}{36} = P(E' \cap G) = P(E') \cdot P(G) = \frac{6}{36} \frac{1}{6}$

$\frac{1}{36} = P(F \cap G) = P(F) \cdot P(G) = \frac{1}{6} \cdot \frac{1}{6}$

But: $\frac{1}{36} = P(E' \cap F \cap G) \stackrel{?}{=} P(E') \cdot P(F) \cdot P(G) = \frac{6}{36} \cdot \frac{1}{6} \cdot \frac{1}{6}$

Kind of makes sense: degrees of freedom

Independence

Definition

Three events E, F, G are *(mutually) independent* if:

$$\begin{aligned}
\mathbf{P}\{E \cap F\} &= \mathbf{P}\{E\} \cdot \mathbf{P}\{F\}, \\
\mathbf{P}\{E \cap G\} &= \mathbf{P}\{E\} \cdot \mathbf{P}\{G\}, \\
\mathbf{P}\{F \cap G\} &= \mathbf{P}\{F\} \cdot \mathbf{P}\{G\}, \\
\mathbf{P}\{E \cap F \cap G\} &= \mathbf{P}\{E\} \cdot \mathbf{P}\{F\} \cdot \mathbf{P}\{G\}.
\end{aligned}$$

say n

For more events the definition is that any (finite) subset of them have this factorisation property.

have to go all the way to n

Examples $0 < p < 1$

n independent events

- every one succeeds

$$\underbrace{p \cdot p \cdots}_{n \text{ times}} = p^n \xrightarrow[n \to \infty]{} 0$$

- at least one succeeds

$$1 - (1-p)^n \xrightarrow[n \to \infty]{} 1$$

- chance that exactly k succeed

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Examples

Fix parameter $0 < p < 1$

n independent experiments, each succeeds with probability p

. chance that every one succeeds?

Examples

Fix parameter $0 < p < 1$

n independent experiments, each succeeds with probability p

- chance that every one succeeds?

$$p^n \xrightarrow[n \to \infty]{} 0$$

- chance that at least one succeeds?

$$1 - (1-p)^n \xrightarrow[n \to \infty]{} 1$$

Murphy's Law

Examples

Fix parameter $0 < p < 1$

n independent experiments, each succeeds with probability p

- chance that every one succeeds?

$$p^n \xrightarrow[n \to \infty]{} 0$$

- chance that at least one succeeds?

$$1 - P(\text{each one fails}) = 1 - (1-p)^n \xrightarrow[n \to \infty]{} 1$$

Murphy's Law

- chance that exactly k succeed?

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Random variables

Random variables

"Random variables $\approx$ random numbers". But *random* means that there must be some kind of experiment behind these numbers.

Definition

A *random variable* is a function from the sample space Ω to the real numbers $\mathbb{R}$.

flip 3 coins X = random variable counting heads

$X: \Omega \to \mathbb{R}$

$X(TTT) = 0$

$X(HTT) = X(THT) = X(TTH) = 1$

$X(HHT) = X(THH) = X(HTH) = 2$

$X(HHH) = 3$

$$P(X=1) = P(X^{-1}(1)) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$$

Random variables

"Random variables $\approx$ random numbers". But *random* means that there must be some kind of experiment behind these numbers.

Definition

A *random variable* is a function from the sample space Ω to the real numbers $\mathbb{R}$.

flip 3 coins X = random variable that counts heads

$X : \Omega \to \mathbb{R}$

$X(TTT) = 0$

$X(HTT) = X(THT) = X(TTH) = 1$

$X(HHT) = X(HTH) = X(THH) = 2$

$X(HHH) = 3$

Random variables

"Random variables $\approx$ random numbers". But *random* means that there must be some kind of experiment behind these numbers.

Definition

A *random variable* is a function from the sample space Ω to the real numbers $\mathbb{R}$.

flip 3 coins X = random variable that counts heads

$X : \Omega \to \mathbb{R}$

$X(TTT) = 0$

$X(HTT) = X(THT) = X(TTH) = 1$

$X(HHT) = X(HTH) = X(THH) = 2$

$X(HHH) = 3$

$P(X=1) = P(X^{-1}(1)) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$

Discrete random variables

Definition

A random variable X that can take on countably many possible values is called *discrete*.

Discrete random variables

Definition

A random variable X that can take on countably many possible values is called *discrete*.

e.g. – nr. heads in 3 coin flips

– nr. of flips needed to first see a head

discrete things I can always count.
don't need to integrate

Probability mass function



Definition

The *probability mass function (pmf)*, or *distribution* of a discrete random variable X gives the probabilities of its possible values:

$$\mathfrak{p}_X(x_i) = \mathbf{P}(X = x_i),$$

Proposition

$$\mathfrak{p}(x_i) \geq 0 \qquad \textit{and} \qquad \sum_i \mathfrak{p}(x_i) = 1$$

Examples

Examples

Fix parameter $\lambda > 0$

Define $p(i) = c \cdot \frac{\lambda^i}{i!}$

which c makes this a pmf.?

$p(i) \geq 0 \iff c \geq 0$

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} c \cdot \frac{\lambda^i}{i!} = c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = c \cdot e^{\lambda} \overset{!}{=} 1 \iff c = e^{-\lambda}$$

Examples

Fix parameter $\lambda > 0$

Define $p(i) = c \cdot \frac{\lambda^i}{i!}$

which c makes this a pmf.?

$p(i) \geq 0 \iff c \geq 0$

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} c \cdot \frac{\lambda^i}{i!} = c \cdot \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = c \cdot e^{\lambda} \iff c = e^{-\lambda}$$

Taylor or Maclaurin Series

Examples

Fix parameter $\lambda > 0$

Define $p(i) = c \cdot \frac{\lambda^i}{i!}$

which c makes this a pmf.?

Taylor or Maclaurin Series

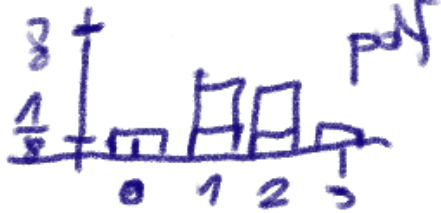
$p(i) \geq 0 \iff c \geq 0$

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} c \cdot \frac{\lambda^i}{i!} = c \cdot \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = c \cdot e^{\lambda} \iff c = e^{-\lambda}$$

so e.g. $P(x=0) = p(0) = e^{-\lambda} \cdot \frac{\lambda^0}{0!} = e^{-\lambda}$

$$\begin{aligned} P(x>2) &= 1 - P(x \leq 2) \\ &= 1 - P(x=0) - P(x=1) - P(x=2) \\ &= 1 - e^{-\lambda} - e^{-\lambda}\lambda - e^{-\lambda}\frac{\lambda^2}{2} \end{aligned}$$

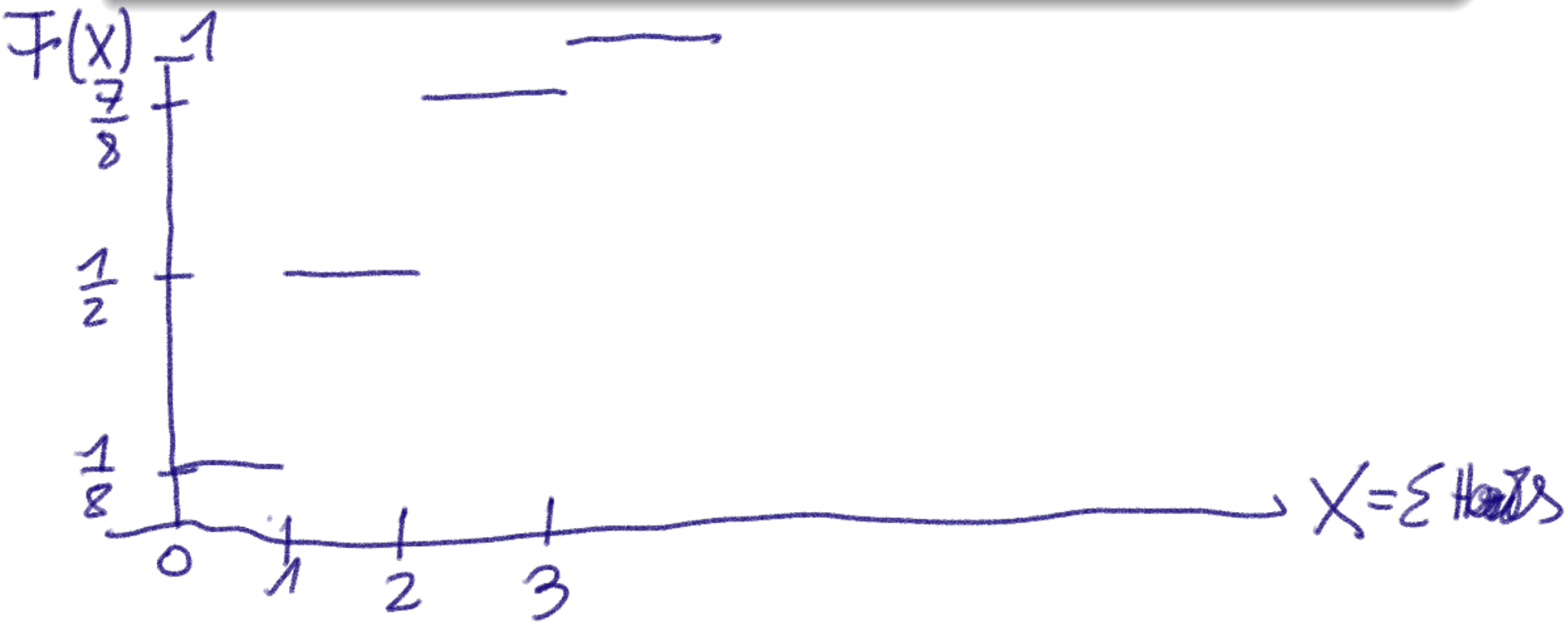
Cumulative distribution function



Definition

The *cumulative distribution function (cdf)* of a random variable X:

$$F : \mathbb{R} \to [0, 1], \qquad x \mapsto F(x) = \mathbf{P}(X \leq x).$$



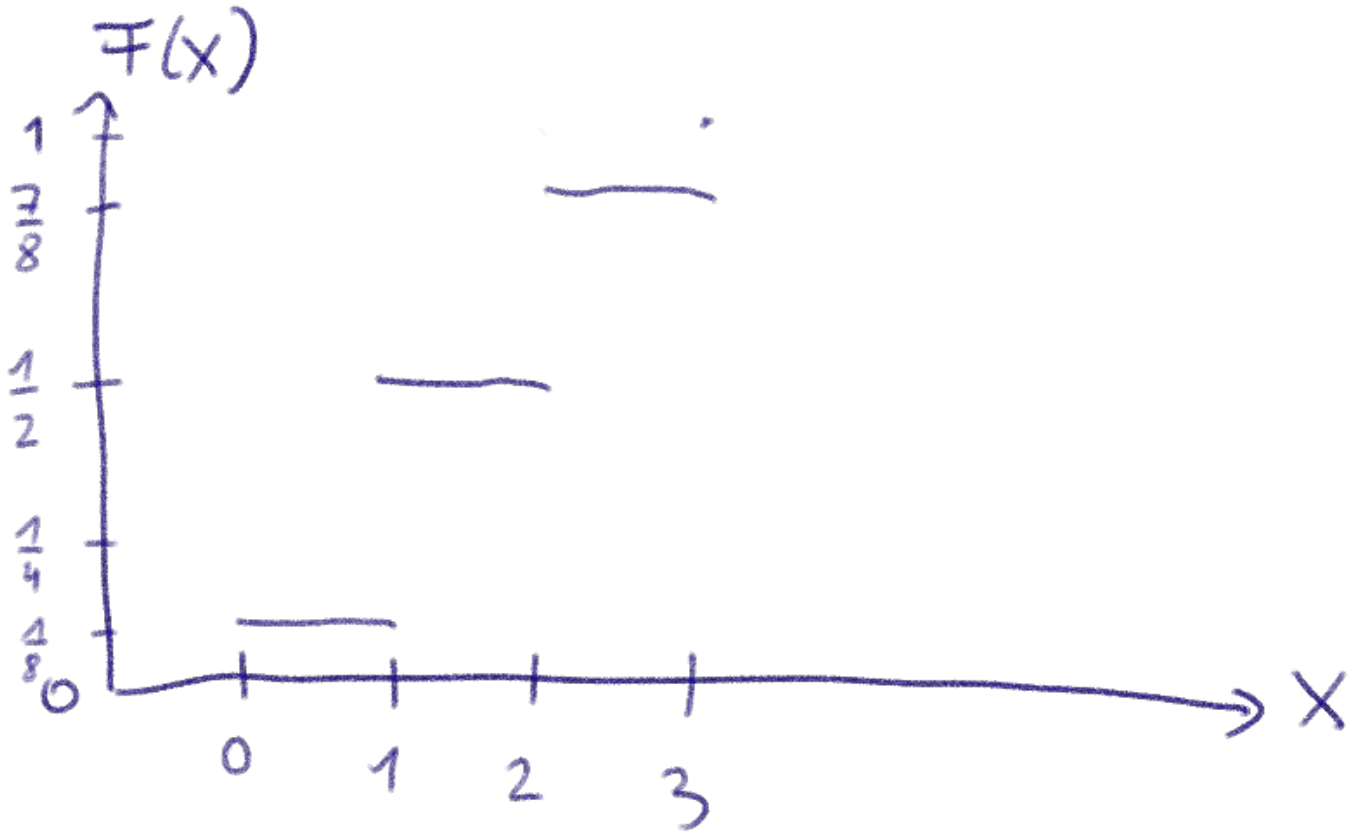
Cumulative distribution function



Definition

The *cumulative distribution function (cdf)* of a random variable X:

$$F : \mathbb{R} \to [0, 1], \qquad x \mapsto F(x) = \mathbf{P}(X \leq x).$$



Cumulative distribution function

Will be very important when working with continuous random variables.

Proposition

A cumulative distribution function F*:*

- *is non-decreasing: if* $x \leq y$ *then* $F(x) \leq F(y)$
- *has limit* $\lim_{x \to -\infty} F(x) = 0$ *on the left*
- *has limit* $\lim_{x \to \infty} F(x) = 1$ *on the right*

Expectation

Expectation

Once we have a random variable, we want to quantify its *typical* behaviour in some sense. Two of the most often used quantities for this are the *expectation* and the *variance*.

Definition

The *expectation* of a discrete random variable X is:

$$\mathbf{E}X = \sum_i x_i \cdot \mathfrak{p}(x_i)$$

provided the sum exists. Also called *mean*, or *expected value*.

weighted average / center of mass

Examples

X = number on a fair die

$$\mathbb{E}X = \sum_i x_i \cdot p(x_i) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} \cdots 6 \frac{1}{6} = \frac{7}{2} = 3.5$$

$$\frac{n \cdot (n+1)}{2} \quad \frac{6 \cdot 7}{2} \frac{1}{6}$$

Examples

X = nr on fair die after roll

$$EX = \sum_{i=1}^{6} x_i \cdot p(X=i) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} \cdots + 6\frac{1}{6} = \frac{7}{2}$$

expectation need not be a possible value

Properties of expectation

Proposition (expectation of a function of a random variable)

If X is a discrete random variable, and $g : \mathbb{R} \to \mathbb{R}$ a function, then:

$$\mathbf{E}g(X) = \sum_i g(x_i) \cdot \mathfrak{p}(x_i) \qquad \textit{(if it exists)}$$

Corollary (expectation is linear)

If X is a discrete random variable, and a, b fixed real numbers:

$$\mathbf{E}(aX + b) = a \cdot \mathbf{E}X + b.$$

Properties of expectation

Proposition (expectation of a function of a random variable)

If X is a discrete random variable, and $g : \mathbb{R} \to \mathbb{R}$ a function, then:

$$\mathbf{E}g(X) = \sum_i g(x_i) \cdot \mathfrak{p}(x_i) \qquad \textit{(if it exists)}$$

Corollary (expectation is linear)

If X is a discrete random variable, and a, b fixed real numbers:

$$\mathbf{E}(aX + b) = a \cdot \mathbf{E}X + b.$$

$$\sum_i c \cdot x_i \cdot p(x_i) = c \cdot \sum_i x_i \cdot p(x_i) \qquad \sum_i c \cdot p(x_i) = c$$

Moments

Definition (moments)

Let $n \in \mathbb{N}$. The n^{th} *moment* of a random variable X is:

$$\mathbf{E}X^n$$

The n^{th} *absolute moment* of X is:

$$\mathbf{E}|X|^n$$

$E x^n = E(X^n)$

Moments

Definition (moments)

Let $n \in \mathbb{N}$. The n^{th} *moment* of a random variable X is:

$$\mathbf{E}X^n$$

The n^{th} *absolute moment* of X is:

$$\mathbf{E}|X|^n$$

$\mathbb{E}X^n = \mathbb{E}(X^n) \neq (\mathbb{E}X)^n$!

Variance

3. Variance

Definition (variance, standard deviation)

The *variance* and the *standard deviation* of a random variable are:

- $\mathbf{Var}\, X = \mathbf{E}(X - \mathbf{E}X)^2$.
- $\mathrm{SD}\, X = \sqrt{\mathbf{Var}\, X}$.

3. Variance

Definition (variance, standard deviation)

The *variance* and the *standard deviation* of a random variable are:

- $\mathbf{Var}\, X = \mathbf{E}(X - \mathbf{E}X)^2$.
- $\mathrm{SD}\, X = \sqrt{\mathbf{Var}\, X}$.

$$\mathrm{Var}\, X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$$

$$\mathbb{E}(X^2 - 2X \cdot \mathbb{E}X + (\mathbb{E}X)^2) = \mathbb{E}(X^2) - 2 \cdot \mathbb{E}X^2 \cdot \mathbb{E}X + (\mathbb{E}X)^2$$

$$= \mathbb{E}X^2 - (\mathbb{E}X)^2$$

Properties of the variance

Proposition (equivalent form of the variance)

$\mathbf{Var}\, X = \mathbf{E}X^2 - (\mathbf{E}X)^2$ *for any random variable* X.

Corollary

$\mathbf{E}X^2 \geq (\mathbf{E}X)^2$ *for any random variable* X,
with equality only if X *is constant.*

Examples

X = roll fair die

$\mathbb{E}X = \frac{7}{2}$

$\mathrm{Var}\, X = \mathbb{E}X^2 - (\mathbb{E}X)^2$

$= 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} \ldots 36 \frac{1}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$

$\mathrm{SD}\, X = \sqrt{\frac{35}{12}} \approx 1.71$

two most important numbers about a fair die: average 3.5

standard deviation 1.7

Topics

- Bernoulli distribution: single trial
- Binomial distribution: many independent trials
- Poisson distribution: counting independent trials
- Geometric distribution: first success in independent trials

Bernoulli and Binomial distributions

~~Bernoulli~~ distribution

Binomial

Definition

Suppose that n independent trials are performed, each succeeding with probability p. Let X count the number of successes within the n trials. Then X has the *Binomial distribution with parameters* n *and* p or, in short, $X \sim Binom(n, p)$.

Special case $n = 1$ is called *Bernoulli distribution with parameter* p.

~~Bernoulli~~ Binomial: mass function

Proposition

Let $X \sim \mathit{Binom}(n, p)$. *Then* $X = 0, 1, \ldots, n$, *and its mass function is*

$$\mathfrak{p}(i) = \mathbf{P}(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \qquad i = 0, 1, \ldots, n.$$

In particular, the $\mathit{Bernoulli}(p)$ *variable can take on values* 0 *or* 1, *with respective probabilities*

$$\mathfrak{p}(0) = 1 - p, \qquad \mathfrak{p}(1) = p.$$

Mass function

$$p(i) = P(X=i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, \dots, n.$$

1) $p(i) \geq 0$

2) $\sum_{i=0}^{n} p(i) = \sum \binom{n}{i} p^i (1-p)^{n-i} = (p + (1-p))^n = 1$

Newton binomial formula

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} \cdot a^i \cdot b^{n-i}$$

$a = p$, $b = 1-p$

Mass function

$$p(i) = P(X=i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, \ldots, n.$$

$p(i) \geq 0$

$$\sum_{i=0}^{n} p(i) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = (p + (1-p))^n = 1$$

Newton binomial formula:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k}$$

Example

Pack of screws of size 10

Prob defective 0.1

If more than one defective screw then pack can be returned

$X \sim \text{Binom}(10, 0.1)$

$$P(X>1) = 1 - P(X=0) - P(X=1)$$
$$- \binom{n}{0} \cdot 0.1^0 \cdot 0.9^n - \binom{n}{1} \cdot 0.1^1 \cdot 0.9^9 = 0.26$$

26

Example

Screws are sold in packs of 10.

Each one is broken with probability 0.1.

If more than one screw is broken, you can return pack.

What percentage of packs is returned

X = nr defective screws

$X \sim \text{Binom}(10, 0.1)$

$$P(X \geq 2) = 1 - P(X=0) - P(X=1)$$
$$= 1 - \binom{10}{0} 0.1^0 \cdot 0.9^{10} - \binom{10}{1} 0.1^1 \cdot 0.9^9 \approx 26\%$$

0.2639

~~Bernoulli~~ Binomial: expectation, variance

Proposition

Let $X \sim Binom(n, p)$. *Then:*

$$\mathbf{E}X = np, \qquad \text{and} \qquad \mathbf{Var}\, X = np(1-p)$$

~~Bernoulli~~ Binomial: expectation, variance

Proposition

Let $X \sim Binom(n, p)$. *Then:*

$$\mathbf{E}X = np, \qquad \text{and} \qquad \mathbf{Var}\, X = np(1-p)$$

Poisson distribution

Poisson: mass function

The Poisson distribution is of central importance in Probability. Will later see relation to Binomial.

Definition

Fix a positive real number λ. The random variable X is *Poisson distributed with parameter* λ, in short $X \sim Poi(\lambda)$, if it is non-negative integer valued, and its mass function is

$$\mathfrak{p}(i) = \mathbf{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}, \qquad i = 0, 1, 2, \ldots$$

ic a p.mf

but why this function

Poisson: mass function

The Poisson distribution is of central importance in Probability. Will later see relation to Binomial.

Definition

Fix a positive real number λ. The random variable X is *Poisson distributed with parameter* λ, in short $X \sim Poi(\lambda)$, if it is non-negative integer valued, and its mass function is

$$\mathfrak{p}(i) = \mathbf{P}(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}, \qquad i = 0, 1, 2, \ldots$$

Indeed a p.m.f used Taylor series to prove this.

Fine, but why this distribution?

Poisson approximation of Binomial

Proposition

Fix $\lambda > 0$, and suppose that $Y_n \sim Binom(n, p)$ with $p = p(n)$ in such a way that $n \cdot p \to \lambda$. Then the distribution of Y_n converges to $Poisson(\lambda)$:

$$n \cdot p(n) \xrightarrow[n \to \infty]{} \lambda$$

$$\forall i \geq 0 \qquad \mathbf{P}(Y_n = i) \underset{n \to \infty}{\longrightarrow} e^{-\lambda} \frac{\lambda^i}{i!}.$$

Poisson approximation of Binomial

Proposition

Fix $\lambda > 0$, and suppose that $Y_n \sim Binom(n, p)$ with $p = p(n)$ in such a way that $n \cdot p \to \lambda$. Then the distribution of Y_n converges to $Poisson(\lambda)$:

$$\forall i \geq 0 \qquad \mathbf{P}(Y_n = i) \underset{n\to\infty}{\longrightarrow} e^{-\lambda}\frac{\lambda^i}{i!}.$$

Take $Y_n = Binom(n,p)$ with large n, small p, such that $np = \lambda$
then Y_n is approximately Poisson distributed with parameter λ.

Proof

$$P(Y_n = i) \longrightarrow e^{-\lambda} \frac{\lambda^i}{i!}$$

$$\binom{n}{i} \cdot (p^i)(1-p)^{n-i}$$

$$\frac{n!}{i!(n-i)!} \quad \frac{1}{i!}(n-1)\cdot(n-2)\cdot(n-3)\cdots(n-i+1) \cdot \frac{(1-p)^n}{(1-p)^i}$$

Proof

$$P(Y_n = i) \xrightarrow[n\to\infty]{} e^{-\lambda} \frac{\lambda^i}{i!}$$

$$P(Y_n = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \frac{1}{i!} \cdot np \cdot (n-1)p \cdot (n-2)p \cdots (n-i+1)p \frac{(1-p)^n}{(1-p)^i}$$

$$\xrightarrow[n\to\infty]{} \frac{1}{i!} \cdot \lambda \cdot \lambda \cdot \lambda \cdots \lambda \cdot e^{-\lambda}$$

$$= \frac{1}{i!} \cdot \lambda^i \cdot e^{-\lambda} = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

$$np \longrightarrow \lambda$$

$$(n-1)p \longrightarrow \lambda$$

$$(n-2)p \longrightarrow \lambda$$

$$(1-p)^n \to \left(1 - \frac{\lambda}{n}\right)^n \longrightarrow e^{-\lambda}$$

$$(1-p)^i \longrightarrow 1$$

Poisson: expectation, variance

Proposition

For $X \sim Poi(\lambda)$, $\mathbf{E}X = \mathbf{Var}\, X = \lambda$.

Example $X \sim \text{Poi}(\lambda)$ $\mathbb{E}X = \text{Var}\, X = \lambda$

Book has average $\frac{1}{2}$ typo per page.

Chance that a random page has ≥ 3 typos?

$X \sim \text{Poisson}(\lambda)$

$\mathbb{E}X = \lambda = \frac{1}{2}$

$P(X \geq 3) = 1 - P(X \leq 2)$

$1 - P(X=0) - P(X=1) - P(X=2)$

$1 - e^{-\frac{1}{2}} \cdot \frac{\frac{1}{2}^0}{0!} - e^{-\frac{1}{2}} \frac{\frac{1}{2}^1}{1!} - e^{-\frac{1}{2}} \frac{\frac{1}{2}^2}{2!} = 1.4\%$

Example $X \sim \text{Poi}(\lambda)$ $\mathbb{E}X = \text{Var}\,X = \lambda$

Book has average $\frac{1}{2}$ typo per page.

Chance that a random page has ≥ 3 typos?

$X \sim \text{Poisson}(\lambda)$

$\mathbb{E}X = \lambda = \frac{1}{2}$

$$P(X \geq 3) = 1 - P(X \leq 2)$$
$$= 1 - P(X=0) - P(X=1) - P(X=2)$$
$$= 1 - e^{-\frac{1}{2}} \cdot \frac{(\frac{1}{2})^0}{0!} - e^{-\frac{1}{2}} \cdot \frac{(\frac{1}{2})^1}{1!} - e^{-\frac{1}{2}} \cdot \frac{(\frac{1}{2})^2}{2!} \approx 1.4\%$$

Example $X \sim \text{Poi}(\lambda)$ $\mathbb{E}X = \text{Var}\,X = \lambda$

Book has average $\frac{1}{2}$ typo per page.

Chance that a random page has ≥ 3 typos?

$X \sim \text{Poisson}(\lambda)$

$\mathbb{E}X = \lambda = \frac{1}{2}$

$$P(X \geq 3) = 1 - P(X \leq 2)$$
$$= 1 - P(X=0) - P(X=1) - P(X=2)$$
$$= 1 - e^{-\frac{1}{2}} \cdot \frac{(\frac{1}{2})^0}{0!} - e^{-\frac{1}{2}} \cdot \frac{(\frac{1}{2})^1}{1!} - e^{-\frac{1}{2}} \cdot \frac{(\frac{1}{2})^2}{2!} \approx 1.4\%$$

Poisson distribution models: many independent trials with small success probability summing up to "a few" in expectation

e.g. – nr of typos on page of book
– nr of citizens over 100 years in big city
– nr of calls per hour in customer center

Examples

Examples

Defective screws.

$X \sim \text{Binom}(10, 0.1)$

Approximated by Poisson(1) as $\lambda = 10 \cdot 0.1 = 1 = np$

$$P(X \geq 2) = 1 - P(X=0) - P(X=1)$$
$$= 1 - e^{-1}\frac{1^0}{0!} - e^{-1}\frac{1^1}{1!} \approx 26\%$$

0.26424

Poisson: History 1



Poisson: History 2



Geometric distribution

Geometric: mass function

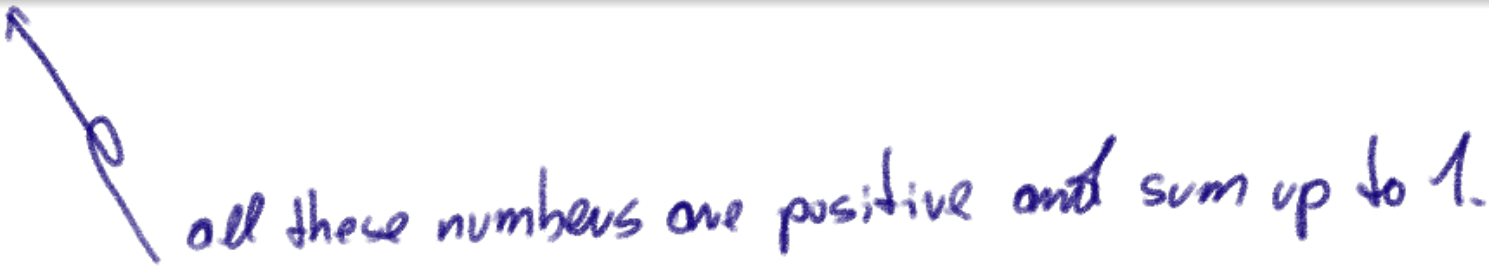
Again independent trials, but now ask: when is the first success?

Definition

Suppose that independent trials, each succeeding with probability p, are repeated until the first success. The total number X of trials made has the *Geometric*(p) distribution (in short, $X \sim Geom(p)$).

Proposition

X can take on positive integers, with probabilities
$\mathbb{p}(i) = (1-p)^{i-1} \cdot p$, $i = 1, 2, \ldots$.



Geometric: mass function

Corollary

The Geometric random variable is (discrete) memoryless:

$$\mathbf{P}\{X \geq n + k \,|\, X > n\} = \mathbf{P}\{X \geq k\}$$

for every $k \geq 1$, $n \geq 0$.

Geometric: expectation, variance

Proposition

two most important numbers

For a Geometric(p) random variable X:

$$\mathbf{E}X = \frac{1}{p} \qquad \mathbf{Var}\, X = \frac{1-p}{p^2}$$

Example

Example

$\mathbb{E}X = \frac{1}{p}$ $\quad \mathrm{Var}\, X = \frac{1-p}{p^2}$

first 3 on fair die takes $X \sim \mathrm{Geom}(\frac{1}{6})$ rolls

$$\mathbb{E}X = \frac{1}{1/6} = 6$$

$$SD(x) = \sqrt{\mathrm{Var}\, X} = \sqrt{\frac{1-1/6}{1/36}} = \sqrt{30} = 5.48$$

Chance that first 3 comes on 7th roll is:

$$P(X=7) = (1-\tfrac{1}{6})^6 \cdot \frac{1}{6} = 5.6\%$$

Summary

- Bernoulli distribution: single trial
- Binomial distribution: many independent trials
- Poisson distribution: counting independent trials
- Geometric distribution: first success in independent trials