PTS Chapter 3 - Final Lecture Notes

3 Special Random Variables

3.1 Bernoulli and Binomial

A Bernoulli experiment has only 2 outcomes

$$X = \begin{cases} 1 & \text{if success} \\ 0 & \text{if failure} \end{cases}$$

Pmf of X:

$$P[X = 1] = p, \quad \text{for some } 0 \leq p \leq 1$$

$$P[X = 0] = (1-p)$$

Let X be a Bernoulli variable with probability p. Then

$$E[X] = 1 \cdot p + 0 \cdot (1-p) = p$$

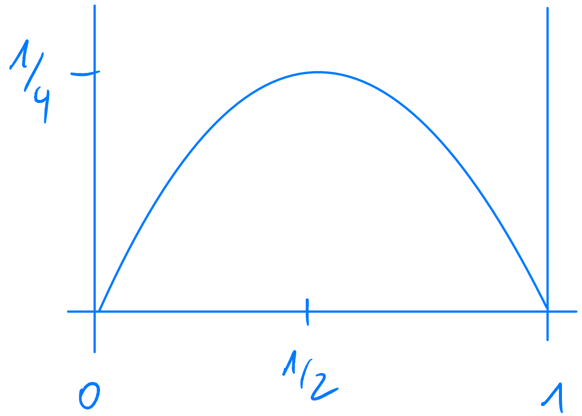
Note: $X^2 = X$

$$E[X^2] = E[X] = p$$

$$Var(X) = E[X^2] - E[X]^2$$

$$= p - p^2 = p(1-p)$$

maximal for $p = \frac{1}{2}$



Graph of $p(1-p)$,

maximal for $p = \frac{1}{2}$

Let us repeat a Bernoulli experiment with RV X.

Suppose $X_1, \dots, X_n, \dots$ are i.i.d. Bern(p) RVs.

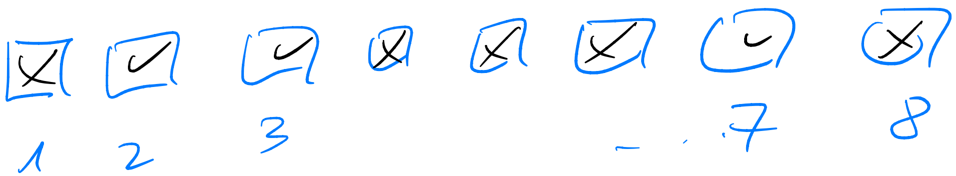
Then

$$Y_n := \sum_{i=1}^{n} X_n$$ counts the number of successes

Possible values of Y_n: $0, 1, \dots, n$

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

Boxes to record success



3 successes

- # ways to get 3 successes? $\binom{8}{3}$
- Probability of such an outcome: $p^3(1-p)^{8-3}$

Let us repeat a Bernoulli experiment with RU X n times

Suppose $X_1, \dots, X_n$ are i.i.d. Bern(p) RUs.

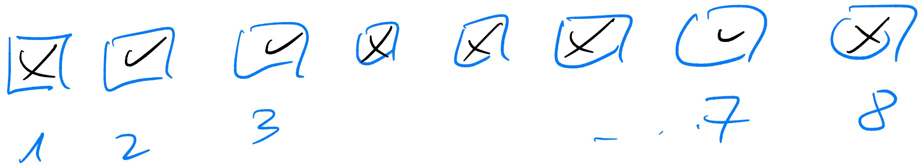
Then

$$Y_n := \sum_{i=1}^{n} X_n$$ counts the number of successes

The possible values of Y_n are $0, 1, 2, \dots, n$.

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

Boxes to record success



3 successes

- # ways to get 3 successes? $\binom{8}{3}$
- Probability of such an outcome: $p^3 (1-p)^{8-3}$

Note:

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

according to the binomial theorem.

Suppose $q = 1-p$. Then

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

This shows that

$$p_i := \binom{n}{i} p^i (1-p)^{n-i}, \quad 0 \le i \le n,$$

is a probability mass function.

We say that Y_n is distributed according to the binomial distribution with parameters n and p, written

$$Y_n \sim B(n, p).$$

We calculate mean and variance:

$$E[Y_n] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

$$Var(Y_n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

$$= \sum_{i=1}^{n} p(1-p) = np(1-p)$$

Example: A satellite system has 4 components and functions if at least 2 are working.

Each component is independently working with probability $p = 0.6$.

What is the probability that the system functions?

$$P[\text{System functions}]$$

$$= 1 - P[\text{System doesn't function}]$$

$$= 1 - p_0$$

$$p_0 = P[\text{all components fail}] + P[\text{exactly 3 components fail}]$$

$$= P[\text{no component functions}] + P[\text{exactly 1 comp. functions}]$$

$$= \binom{4}{0} p^0 (1-p)^4 + \binom{4}{1} p^1 (1-p)^3$$

$$= 1 \cdot 0.4^4 + 4 \cdot 0.6 \cdot 0.4^3 = 0.1792$$

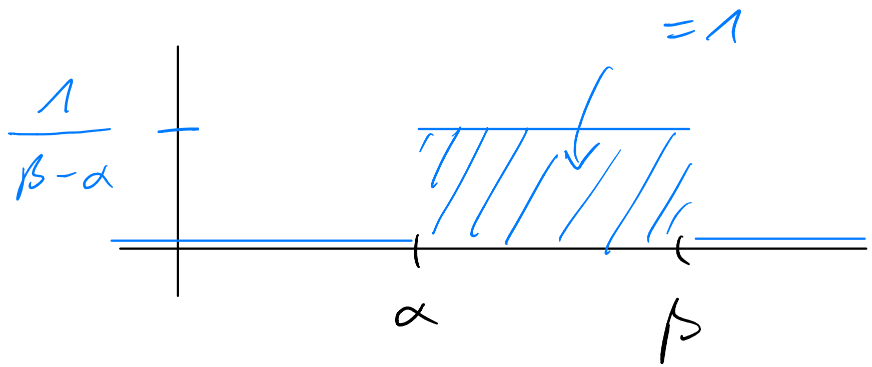
3.3 Uniform Random Variables

A continuous RV X is uniformly distributed if there is an interval $[\alpha, \beta]$ so that

- X takes only values in $[\alpha, \beta]$
- all values are equally probable.

This means, X has the density f where

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, & \alpha \leq x \leq \beta \\ 0 & \text{otherwise.} \end{cases}$$



We write $X \sim U[\alpha, \beta]$.

What about mean and variance?

We first determine mean and variance for the simple case that $X \sim U[0,1]$. Then

$$E[X] = \int_0^1 x \cdot 1 \, dx = \left[\frac{x^2}{2} \right]_0^1 = \frac{1}{2}$$

$$E[X^2] = \int_0^1 x^2 \cdot 1 \, dx = \left[\frac{x^3}{3} \right]_0^1 = \frac{1}{3}$$

Hence,

$$Var(X) = E[X^2] - E[X]^2$$

$$= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{4}$$

Suppose now that $Y \sim U[\alpha, \beta]$.

Then $X := \dfrac{Y-\alpha}{\beta-\alpha}$ is $U[0,1]$-distributed and

$$Y = (\beta-\alpha) X + \alpha.$$

Therefore,

$$E[Y] = (\beta-\alpha) E[X] + \alpha = \frac{\beta-\alpha}{2} + \alpha$$

$$= \frac{\alpha+\beta}{2}$$

$$Var(Y) = (\beta-\alpha)^2 Var(X) = \frac{(\beta-\alpha)^2}{12}$$

Mean of $U[\alpha, \beta]$: Brute-Force Calculation

This is a calculation of mean and variance following directly the definition. Compare this to our approach of (i) solving a simple variant of the problem and (ii) reducing complex cases to the simple one.

Also:

$$E[X] = \int_\alpha^\beta x \frac{1}{\beta-\alpha} dx = \frac{1}{\beta-\alpha} \left[\frac{x^2}{2} \right]_\alpha^\beta$$

$$= \frac{1}{\beta-\alpha} \frac{1}{2} (\beta^2 - \alpha^2) = \frac{(\beta+\alpha)(\beta-\alpha)}{2(\beta-\alpha)}$$

$$= \frac{\alpha+\beta}{2}$$

Variance of $U[\alpha,\beta]$: Brute-Force Calculation

$$E[X^2] = \int_\alpha^\beta x^2 \frac{1}{\beta-\alpha}\, dx = \frac{1}{\beta-\alpha}\left[\frac{x^3}{3}\right]_\alpha^\beta$$

$$= \frac{1}{3}\,\frac{\beta^3-\alpha^3}{\beta-\alpha} = \frac{1}{3}\left(\beta^2+\alpha\beta+\alpha^2\right)$$

$$Var(X) = E[X^2]-E[X]^2$$

$$= \frac{\beta^2+\alpha\beta+\alpha^2}{3} - \frac{(\alpha+\beta)^2}{2^2}$$

$$= \frac{4\beta^2+4\alpha\beta+4\alpha^2-3\alpha^2-6\alpha\beta-3\beta^2}{12}$$

$$= \frac{1}{12}\left(\beta^2-2\alpha\beta+\alpha^2\right) = \frac{(\beta-\alpha)^2}{12}$$

Exponential Functions

Three concepts: We consider functions with the following properties:

1) $f(x) = a^x$, $x \in \mathbb{Q}$ (i.e., x like $\frac{u}{u}$, $\frac{13}{9}$, ...), for some $a > 0$

Exponentiation

$a^1, a^2, \quad a^{1/4} = \sqrt[4]{a}, \quad a^0 = 1$

$1 = a^{-3} \cdot a^3 = a^0 \implies a^{-3} = \frac{1}{a^3}$

$a^{-\frac{7}{3}} = \frac{1}{\sqrt[3]{a^7}}$

2) $f(x+y) = f(x) \cdot f(y)$, $x, y \in \mathbb{R}$

Addition-multiplication homomorphism

3) $f'(x) = \alpha f(x)$, $x \in \mathbb{R}$ and $f(0) = 1$, for some $\alpha \neq 0$.

Growth proportional to value

like

$(e^{\alpha x})' = \alpha e^{\alpha x}$

We will see that these three concepts are actually equivalent for differentiable functions.
If a differentiable function f has one of these three properties, then it also has the other two.

We show that

1) implies 2)

2) implies 3)

3) implies 2)

2) implies 1)

We note that exponentiation can also be defined for real numbers as exponents. This, however, is only conceptually interesting, it does not give us a practical way to compute such powers. That will come later.

If $x \in \mathbb{R}$ is a real number, then we can approximate it by rational numbers. That is, there is a sequence r_n such that

$$\lim_{n \to \infty} r_n = x, \quad \text{or} \quad r_n \to x.$$

Then we define

$$a^x := \lim_{n \to \infty} a^{r_n}$$

For instance, this gives

$$5^\pi = \lim \left(5^3, 5^{\frac{31}{10}}, 5^{\frac{314}{100}}, \dots \right)$$

Implication 1) $\Rightarrow$ 2)

If $f: \mathbb{R} \to \mathbb{R}$ is differentiable and for some $a \in \mathbb{R}$

$$f(x) = a^x, \qquad x \in \mathbb{Q}$$

Exponentiation

then

$$f(x+y) = f(x) \cdot f(y), \quad x, y \in \mathbb{R}$$

Addition – multiplication homomorphism

Proof: By the properties of exponentiation, we have

$$f(x+y) = f(x) \cdot f(y) \qquad \text{for all } x, y \in \mathbb{Q}.$$

If f is differentiable, it is also continuous and the second equation holds also for $x, y \in \mathbb{R}$ because addition and multiplication are continuous.

Implication 2) $\Rightarrow$ 3)

If $f: \mathbb{R} \to \mathbb{R}$ is differentiable and

$$f(x+y) = f(x) \cdot f(y), \quad x, y \in \mathbb{R}$$

then there is a constant $a \in \mathbb{R}$ such that

$$f'(x) = a f(x), \quad x \in \mathbb{R} \qquad \text{and} \qquad f(0) = 1.$$

Proof: First, we note that $f(0) = 1$.

This is because

$$f(0) = f(0+0) = f(0) \cdot f(0)$$

$$\Rightarrow \quad 1 = f(0)$$

Next, we see what we can conclude about f':

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

$$= \lim_{h \to 0} \frac{f(x) \cdot f(h) - f(x) \cdot 1}{h}$$

$$= \lim_{h \to 0} \left(\frac{f(h) - 1}{h} \cdot f(x) \right)$$

$$= \left(\lim_{h \to 0} \frac{f(h) - f(0)}{h} \right) \cdot f(x)$$

$$= f'(0) \cdot f(x)$$

So, $f'(0)$ is the α we were looking for.

Implication 3) $\Rightarrow$ 2)

If $f: \mathbb{R} \to \mathbb{R}$ is differentiable and there is a constant $\alpha \in \mathbb{R}$ such that

$$f'(x) = \alpha f(x), \quad x \in \mathbb{R} \qquad \text{and} \qquad f(0) = 1$$

then

$$f(x+y) = f(x) \cdot f(y), \quad x, y \in \mathbb{R}$$

Proof: This argument is a bit lengthy. We first check that it is enough to prove the claim for $\alpha = 1$. From there we arrive at the power series of the exponential function and obtain the homomorphism equation.

First: It's enough to consider $\alpha = 1$.

Suppose that $g'(x) = \alpha\, g(x)$ and $g(0) = 1$.

We normalize g as f, defined as $f(x) := g\left(\frac{1}{\alpha}x\right)$.

We get back g from f because $g(x) = g\left(\frac{1}{\alpha}\cdot \alpha \cdot x\right) = f(\alpha x)$.

Then $f'(x) \overset{\text{chain rule}}{=} g'\left(\frac{1}{\alpha}x\right)\cdot\frac{1}{\alpha} \overset{\text{proportional growth}}{=} \cancel{\alpha}\, g\left(\frac{1}{\alpha}x\right)\cdot\cancel{\frac{1}{\alpha}} = g\left(\frac{1}{\alpha}x\right) = f(x)$.

That is, f satisfies $f' = f$.

We also have $f(0) = g\left(\frac{1}{\alpha}0\right) = g(0) = 1$.

Suppose we can show that such an f also satisfies

$$f(x+y) = f(x)\cdot f(y).$$

Then $g(x+y) = f(\alpha(x+y)) = f(\alpha x + \alpha y) = f(\alpha x)\cdot f(\alpha y) = g(x)\cdot g(y)$.

So, it suffices to consider $\alpha = 1$.

Second: What does f look like?

It cannot be a polynomial like $f(x) = a_0 + a_1 x^1 + \dots + a_n x^n$.
This would yield $f^{(n+1)} = 0$. $f^{(n+1)}$ is the $(n+1)$-th derivative

Let us assume that f has a power series, i.e., f is an infinitely long polynomial $f(x) = a_0 + a_1 x^1 + \dots + a_n x^n + \dots$.
What does that tell us about the coefficients a_n?

$$f(x) = a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \dots + a_n x^n + a_{n+1} x^{n+1}$$

$$f'(x) = 0 + a_1 x^0 + 2 a_2 x^1 + 3 a_3 x^2 + \dots + n \cdot a_n x^{n-1} + (n+1) a_{n+1} x^n$$

The series f' and f are the same iff they have the same coefficients: $a_1 = a_0$, $2 a_2 = a_1$, $\dots$, $(n+1) a_{n+1} = a_n$.

That is, they satisfy the recurrence

$$a_{n+1} = \frac{1}{n+1} a_n \quad \text{with} \quad a_0 = 1.$$

The recurrence

$$a_{n+1} = \frac{a_n}{n+1} \qquad \text{with} \qquad a_0 = 1$$

leads to the values

$$a_0 = 1, \qquad a_1 = \frac{1}{1}, \qquad a_2 = \frac{1}{1 \cdot 2}, \qquad a_3 = \frac{1}{1 \cdot 2 \cdot 3}$$

and generally

$$a_n = \frac{1}{n!}.$$

The shape of f is therefore

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

This function is also known as the exponential function and it is often denoted as exp.

We see, its form derives from the two conditions

$$f' = f \qquad \text{and} \qquad f(0) = 1.$$

Third: The exponential function satisfies $f(x+y) = f(x) \cdot f(y)$.

We start with the right-hand side:

$$f(x) \cdot f(y) = \left(\sum_{i=0}^{\infty} \frac{x^i}{i!} \right) \cdot \left(\sum_{j=0}^{\infty} \frac{y^j}{j!} \right)$$

Reorganize the sum, combine factors whose exponents add up to n, for each n.

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!}$$

Multiply the inner sum by $1 = \frac{n!}{n!}$.

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k y^{n-k}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Binomial formula!

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (x+y)^n = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = f(x+y)$$

Implication 2) $\Rightarrow$ 1)

If

$$f(x+y) = f(x) \cdot f(y), \quad x, y \in \mathbb{R}$$

then for some $a \in \mathbb{R}$

$$f(x) = a^x, \qquad x \in \mathbb{Q}$$

Proof: We have already shown that from our assumption it follows that

- $f(0) = 1.$

We conclude from $1 = f(0) = f(x+(-x)) = f(x) \cdot f(-x)$ that

- $f(-x) = \dfrac{1}{f(x)}.$

Moreover,

- $$f(m\cdot x) = f(\underbrace{x+\cdots+x}_{m \text{ times}}) = \underbrace{f(x)\cdot\cdots\cdot f(x)}_{m \text{ times}} = f(x)^m$$

From $f(x) = f(\underbrace{\tfrac{x}{n}+\tfrac{x}{n}+\cdots+\tfrac{x}{n}}_{n \text{ times}}) = f(\tfrac{x}{n})^n$

we conclude

- $$f\left(\frac{x}{n}\right) = \sqrt[n]{f(x)} = f(x)^{1/n}$$

Hence, for every rational number $\frac{m}{n}$ we have

- $$f\left(\frac{m}{n}\right) = f\left(m\cdot\frac{1}{n}\right) = f\left(\frac{1}{n}\right)^m = \left(f(1)^{1/n}\right)^m = f(1)^{m/n}$$

So far we have seen that $f(x+y) = f(x) \cdot f(y)$

implies $$f(x) = f(0)^x, \quad x \in \mathbb{Q}$$

For the special case of $f = \exp$, that is, $f' = f$,

we have $$f(1) = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$$

We often denote the number $\exp(1)$ simply as e.

Then we have
$$e^x = \exp(1)^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{Q}$$

Since exp is differentiable, (this was always our assumption)

it is continuous on $\mathbb{Q}$, this equality also holds for $x \in \mathbb{R}$.

If g satisfies $g'(x) = \alpha g(x)$, then $g(x) = \exp(\alpha x)$,
as seen before, that is,

$$g(x) = \sum_{n=0}^{\infty} \frac{\alpha^n x^n}{n!}.$$

We have $\exp(x) > 0$ for $x > 0$ and $\exp(-x) = \frac{1}{\exp(x)}$,
hence $\exp(x) > 0$ holds also for $x < 0$.

That is, $\exp'(x) = \exp(x) > 0$ for all $x \in \mathbb{R}$.

Thus, exp is strictly monotonic and has an inverse function that we call log.

As the inverse of exp, the function log inherits the property

$$\log(x \cdot y) = \log(x) + \log(y).$$

The known laws for logarithms and exponents can all be derived from the development shown so far.

3.4 Exponential Random Variables

Let $\mathcal{X}$ be a random variable that stands for the time we have to wait for a radioactive atom to decay (or for some other similar event).

We assume that the waiting time does not depend on the time we have already waited. To some extent this holds also when waiting for

- the next customer
- the next email
- the next taxi.

Let $\mathcal{X}$ be the waiting time. Our assumption that the probability distribution for the time after s, if the event has not occured, is the same as the original distribution can be expressed as

$$(*) \quad P[\mathcal{X} > s+t \mid \mathcal{X} > s] = P[\mathcal{X} > t]$$

let $F(t) := P[\mathcal{X} \leq t]$ and $G(t) := P[\mathcal{X} > t] = 1 - F(t)$.

Then

$$\lim_{t \to \infty} G(t) = 0$$

because $G(t) = 1 - F(t)$ and $\lim_{t \to \infty} F(t) = 1$.

The definition of conditional probabilities tells us that

(*) is equivalent to

$$\frac{P[X > s+t]}{P[X>s]} = \frac{P[X > s+t, \; X > s]}{P[X>s]}$$

$$= P[X > s+t \mid X > s] = P[X>t]$$

That is,

$$G(s+t) = P[X > s+t]$$

$$= P[X>s] \cdot P[X>t] = G(s) \cdot G(t)$$

This yields

$$G(t) = a^t \quad \text{where} \quad a = G(1)$$

and $a < 1$ since $\lim_{t\to\infty} G(t) = 0$.

Since $a < 1$, we have $\log a < 0$.

Let $\lambda := -\log a$. Then $a^t = e^{(\log a)t} = e^{-\lambda t}$.

Hence, $G(t) = P[T > t] = e^{-\lambda t}$.

$$\Rightarrow \quad F(t) = P[T \leq t] = 1 - P[T > t] = 1 - G(t)$$

$$= 1 - e^{-\lambda t}$$

is the cdf (= distribution function) of T

$$\Rightarrow \quad f(t) = \frac{d}{dt}\, 1 - e^{-\lambda t} = \lambda e^{-\lambda t}$$ is the pdf of T.

We say T is exponentially distributed with parameter λ, written

$$T \sim \mathrm{Exp}(\lambda)$$

What does λ stand for? The dimension of t is time.

$\Rightarrow$ The dimension of λ is time^{-1},

i.e., λ is a frequency or rate.

In the labs we have calculated:

(using integration by parts)

- $$E[X] = \int_0^\infty t \cdot e^{-\lambda t}\, dt = \frac{1}{\lambda}$$

$\frac{1}{\lambda}$ is the average waiting time

$$E[X^2] = \int_0^\infty t^2 \cdot e^{-\lambda t}\, dt = \frac{2}{\lambda^2}$$

λ is the average number of events per time unit, i.e., the rate of events

- $$Var(X) = E[X^2] - E[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Hence, $\mu = \frac{1}{\lambda}$, $\sigma = \frac{1}{\lambda}$

Multiple Mailboxes

We assume that the arrival of E-mails can be modeled by an exponential distribution. That is, there is a rate $\lambda > 0$ such that the probability to wait at least for a time t for the next mail is

$$e^{-\lambda t} = G(t)$$

Suppose there are n people with an E-mail mailbox and the rate at which mail arrives at maibox i is λ_i.

What is the probability that no message will arrive at any of the mailboxes during the next time period of t if arrivals at different boxes are independent?

Let X_i be the waiting time for a message to arrive at mailbox i.

Then $X_i \sim Exp(\lambda_i)$.

The probability that no mail arrives at box i during time t is

$$P[X_i > t] = e^{-\lambda_i t} \qquad = G_i(t)$$

The probability that no mail arrives at any box is

$$P[X_1 > t \,\&\, \dots \,\&\, X_n > t]$$

$$= P[X_1 > t] \cdot P[X_2 > t] \cdot \ldots \cdot P[X_n > t] \qquad \text{independence of } X_i$$

$$= e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} \cdot \ldots \cdot e^{-\lambda_n t} \qquad G_1(t) \cdots G_n(t)$$

$$= e^{-(\lambda_1 + \lambda_2 + \dots + \lambda_n)t}$$

<u>Proposition 60</u>: If $X_1, \dots, X_n$ are independent RVs, $X_i \sim Exp(\lambda_i)$, then

$$\min(X_1, \dots, X_n) \sim Exp(\lambda_1 + \dots + \lambda_n)$$

3.2 The Poisson Distribution

The Poisson distribution models a scenario where a sequence of events happens:

- the time between events is distributed exponentially with rate λ
- the times between two events are independent of each other.

We are then interested in how many events happen during an interval of unit length (the length to which the rate λ refers.

The Poisson distribution gives us the probability that exactly k events happen during a unit interval.

To apply it, we need the rate λ and we have to verify that the underlying assumptions hold.

Examples: Essentially the same as for the exponential distribution. Question answered can be about the number of

- customers arriving per hour at a shop during a workday afternoon
- emails arriving per minute at a mailserver
- soldiers being killed by horse kicks per year (classical application in Germany around 1900)

Whether assumption are satisfied can be checked by

- measuring the average waiting time T
- checking whether the times are $\mathrm{Exp}\left(\frac{1}{T}\right)$ distributed.

Let $X_1, X_2, \ldots$ be independent exponentially distributed RVs with rate λ. We interpret the X_i as consecutive waiting times:

- X_1 is the time until the first event happens
- X_2 is the subsequent time until the second event happens

etc.

What is the probability that exactly k events happen during the interval $[0,1]$

(e.g., within one hour, one day etc.)?

This problem deals with the sum of i.i.d $\mathrm{Exp}(\lambda)$ RVs.

Given the $\mathcal{X}_i$, let

$$S_k := \sum_{i=1}^{k} \mathcal{X}_i .$$

sum of waiting times for first k events

and let

$$N := \arg\max_k \left(S_k \leq 1 \right),$$

maximal k s.th. k events happen in one time unit

that is, N is the maximal number of consecutive $\mathcal{X}_i$, starting with $i=1$, whose sum does not exceed 1. Note, N is discrete. What is

$$P[N=k],$$

$k = 0, \ldots, k, \ldots$?

Probability of exactly k events in a unit time

Proposition:

$$N = k \iff S_k \le 1 \text{ and } X_{k+1} \ge 1 - S_k$$

Plan: Let f be the density of X_{k+1} and f_k of S_k.

Then

- $f(s) \cdot f_k(t)$ is the joint distribution of X_{k+1} and S_k

- $P[N = k] = P[S_k \le 1, X_{k+1} > 1 - S_k]$

$$= \boxed{\int_0^1 f_k(t) \int_{1-t}^{\infty} f(s)\, ds\, dt}$$

Note:
- $S_k \ge 0$
- $X_k \ge 0$
- S_k, X_{k+1} indep

We know $f(s) = \lambda e^{-s}$.

But what is f_k?

General consideration: Let X, Y be independent,

$X \sim f(x)$, $Y \sim g(y)$.

Then

$$X + Y \sim f * g$$

where

$$\boxed{(f*g)(z) = \int_{-\infty}^{\infty} f(x) \cdot g(z-x)\, dx}$$

- $f*g$ is the convolution of f and g

- Iterate over all combinations of numbers that sum up to z:

  $x + (z-x) = z$

- Multiply their probabilities:

  $f(x) \cdot g(z-x)$

- Sum them up:

  integrate

Find out f_k!

- $f_1(t) = \lambda e^{-\lambda t}$

- $f_2(t) = (f_1 * f_1)(t) = \int_0^t f_1(s) f_1(t-s)\, ds$

$$= \int_0^t \lambda e^{-\lambda s} \cdot \lambda e^{-\lambda(t-s)}\, ds$$

$$= \lambda^2 \int_0^t e^{-\lambda(s+t-s)}\, ds$$

$$= \lambda^2 \int_0^t e^{-\lambda t}\, ds = \lambda^2 e^{-\lambda t} \int_0^t ds$$

$$= \lambda^2 t e^{-\lambda t}$$

- $$f_3(t) = (f_2 * f_1)(t)$$
$$= \int_0^t \lambda^2 s e^{-\lambda s} \lambda e^{-\lambda(t-s)} ds$$
$$= \lambda^3 e^{-\lambda t} \int_0^t s\, ds = \lambda^3 \frac{t^2}{2} \cdot e^{-\lambda t}$$

- $$f_k(t) = \lambda^k \frac{t^{k-1}}{(k-1)!} e^{-\lambda t} \sim S_k$$

Density of k-fold sum of i.i.d exponential RV.s with frequency λ

Also called Gamma distribution $\Gamma(k, \frac{1}{\lambda})$

Now, let's calculate:

$$P[N=k] = P[\overset{\sim f_k}{S_k \le 1}, \overset{\sim f}{X_{k+1} > 1 - S_k}]$$

$$= \int_0^1 f_k(t) \left(\int_{1-t}^{\infty} f(s)\, ds \right) dt$$

$$= \int_0^1 \lambda^k \frac{t^{k-1}}{(k-1)!} e^{-\lambda t} \int_{1-t}^{\infty} \lambda e^{-\lambda s}\, ds \, dt$$

$$= \frac{\lambda^k}{(k-1)!} \int_0^1 t^{k-1} e^{-\lambda t} \left[-e^{-\lambda s} \right]_{1-t}^{\infty} dt$$

$$= \frac{\lambda^k}{(k-1)!} \int_0^1 t^{k-1} e^{-\lambda t} e^{-\lambda(1-t)} dt$$

We continue:

$$P[N=k] = \frac{\lambda^k}{(k-1)!} \int_0^1 t^{k-1} e^{-\lambda t} \; e^{-\lambda(1-t)} \, dt$$

$$= \frac{\lambda^k}{(k-1)!} \int_0^1 t^{k-1} e^{-\lambda} \, dt$$

$$= \frac{\lambda^k}{(k-1)!} \int_0^1 t^{k-1} \, dt \; e^{-\lambda}$$

$$= \frac{\lambda^k}{(k-1)!} \left[\frac{t^k}{k} \right]_0^1 e^{-\lambda} = \frac{\lambda^k}{k!} e^{-\lambda}$$

This is the pmf of the Poisson distribution with rate λ, $\mathrm{Pois}(\lambda)$

Example 56: Assume, on average there are three accidents per week on the highway between Trento and BZ. What is the probability that there is at least one accident this week?

$\lambda = 3$

Three accidents per week $\Rightarrow$ frequency $\lambda = 3$

A = # accidents $\sim$ Pois(3)

In general: $P[m \leq A \leq n] = \sum_{k=m}^{n} P[A = k] = \frac{\lambda^k}{k!} e^{-\lambda}$

Here: $P[A \geq 1] = 1 - P[A \leq 0]$

$= 1 - P[A = 0]$

$= 1 - \frac{3^0}{0!} e^{-3} = 1 - e^{-3}$

Probability of at least 5 accidents per week.

$$P[\mathcal{A} \geq 5] = \sum_{k=5}^{\infty} \frac{3^k}{k!} e^{-3}$$

$$= 1 - \sum_{k=0}^{4} \frac{3^k}{k!} e^{-3} = 1 - P[\mathcal{A} \leq 4]$$

Probability of at least 5 accidents in two weeks:

- new unit time: 2 weeks instead of 1
- new frequency: 6 per two weeks
- new RV $\mathcal{A}_2$ (= # accidents in 2 weeks)
  $\sim \text{Pois}(3+3) = \text{Pois}(6)$

$$\Longrightarrow P[\mathcal{A}_2 \geq 5] = 1 - \sum_{k=0}^{4} \frac{6^k}{k!} e^{-6} = 1 - P[\mathcal{A}_2 \leq 4]$$

Mean and variance

Exp(λ) has rate λ,

i.e., λ events per time unit

$\Longrightarrow$ Pois(λ) has mean λ ?

Proof: Let $X \sim \text{Pois}(\lambda)$. Then

$$E[X] = \sum_{k=\cancel{0}\,1}^{\infty} \cancel{k} \frac{\lambda^k}{\cancel{k}!_{\,(k-1)!}} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda}$$

$$= \lambda e^{\lambda} \cdot e^{-\lambda} = \lambda$$

$$\boxed{E[X^2]} = \sum_{k=\cancel{0}\,1}^{\infty} k^{\cancel{2}} \frac{\lambda^k}{\underset{(k-1)!}{\cancel{k}!}} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!}$$

$$= \lambda e^{-\lambda} \sum_{k=0}^{\infty} (k+1) \frac{\lambda^k}{k!}$$

$$= \lambda e^{-\lambda} \left(\sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} 1 \cdot \frac{\lambda^k}{k!} \right)$$

$$= \lambda e^{-\lambda} (\lambda e^{\lambda} + e^{\lambda}) = \lambda^2 + \lambda$$

$$Var(X) = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

So, $\mu = \lambda$ and $\sigma^2 = \lambda$

Poisson and Binomial

Suppose $X \sim B(n,p)$

$$P[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

$$= \frac{n!}{k!\,(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

Mean of $B(n,p) = n \cdot p$

Mean of $Pois(\lambda) = \lambda$

View $\lambda = n \cdot p$ *)

$\Rightarrow p = \frac{\lambda}{n}$

*) Idea: Probability p (small!!) for a car to have an accident.

Many cars, n (large!!).

$\Longrightarrow$ Rate of accidents $= n \cdot p = \lambda$.

$$P[X=k] = \frac{n!}{k!\,(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$$

$$= \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\approx 1}\;\frac{\lambda^k}{k!}\;\frac{\overbrace{\left(1-\frac{\lambda}{n}\right)^n}^{\to\, e^{-\lambda}}}{\underbrace{\left(1-\frac{\lambda}{n}\right)^k}_{\to\, 1}}$$

$$\lim_{n\to\infty}\frac{\lambda}{n} = 0$$

Therefore, $P[X=k] \approx \frac{\lambda^k}{k!}e^{-\lambda}$

or

$$B(n,p) \approx \mathrm{Pois}(n\cdot p)$$

for large n and small p.

Note: This is a rule of thumb from the time when computers were rare and slow.

Example 58

The number of customers in a bar is on average 4 per hour. What is the probability that there are no more than 3 in 2 hours?

$\mathcal{C}$: # customers / hour $\sim \text{Pois}(4)$

$\mathcal{C}_2$: # customers / 2 hours $\sim \text{Pois}(4+4)$

$= \text{Pois}(8)$

$P[\mathcal{C}_2 \leq 3]$

$\mathcal{C}_{\text{hour } 1} + \mathcal{C}_{\text{hour } 2}$

indep.

Example 58

The number of customers in a bar is on average 4 per hour. What is the probability that there are no more than 3 in 2 hours?

Let $\mathcal{X}_1$ = # customers in 1^{st} hour

$\mathcal{X}_2$ = # customers in 2^{nd} hour

Remember the Poisson story

- $\mathcal{X}_1, \mathcal{X}_2$ independent $\Rightarrow$ $\mathcal{X}_1 + \mathcal{X}_2$ Poisson
- Rate of $\mathcal{X}_1 + \mathcal{X}_2$ is $4+4=8$ in 2 hours

Reproductive property of the Poisson

$$P[\mathcal{X}_1 + \mathcal{X}_2 \leq 3] = \sum_{k=0}^{3} e^{-8} \frac{8^k}{k!} = 0.423$$

The Poisson distribution is reproductive in the following sense.

Proposition: Let $X_1 \sim Pois(\lambda_1)$, $X_2 \sim Pois(\lambda_2)$, X_1, X_2 ind.

Then

$$X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$$

Proof (by story): X_1 counts events of type 1 with rate λ_1, X_2 events of type 2 with rate λ_2, which happen independently. What is the rate at which both kinds of events happen? Clearly, $\lambda_1 + \lambda_2$.

Alternative proofs by calculation (see lecture notes of 19/20 or scriptum.

Suppose there is a shop visited by λ costumers per hour.

Suppose that a fraction of p are female and of $(1-p)$ are male.

How is the number of female customers distributed?

F # female customers per hour

$F \sim \text{Pois}(p\lambda)$

Suppose there is a shop visited by λ costumers per hour.

Suppose that a fraction of p are female and of $(1-p)$ are male.

How is the number of female customers distributed?

Answer: $\mathrm{Pois}(p\lambda)$, since $p\lambda$ is their rate of arrival.