Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:43

(Real-world)

More Distributions

which disb to use where?

Poisson
Exponential
Geometric
Log-Normal

100's of disb

Confidence Intervals (DS/MLE)

Bootstrapping

CLT

Next-Week: T, T, S

1 / 68

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:43

Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe | New Tab

Search Google or type a URL

Update

Gmail Images

Google

Search Google or type a URL

My Drive
Learning
InterviewBit S...
GitHub
YouTube
Colaboratory
Scaler Acade...
Colaboratory
Google Maps
Add shortcut

Photo by NASA Image Library

2 / 69

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:15

which disb to use when?

① conditions for a disb to arise

↳ real-world case satisfies the conditions

(examples)

Binomial →

- n independent bernoulli trails,
- p : prob of success in each trail

cnt of successes

3 / 6

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:09

② $X : x_1, x_2, x_3 \ldots x_n$ independent obs

e.g: heights; baby weights

(ks-test) → QQ-plot

Statistical Test/plots to determine

if $X \sim$ dist

[Gaussian]

4 / 4

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:11

t-test → hypothesis testing

↳ t-distr (next 2 classes)

5 / 5

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:45

Poisson Dist

poisson r.v

$X =$ number of events per unit time

(e.g) # visitors to a website in 1hr (Amazon.com)

Conditions:
1. constant rate (λ) events
2. events must be indep
3. no simultaneous events (μsec gap)

poisson process

(e.g) # ~~visitors~~ groups to a restaurent in 1hr during weekends b/w 7-9PM

1. Constant rate (λ) of events → ✓

2. events are indep → not fully true

3. no simultaneous events → family/friends ✗

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:24

discrete
or
continous

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

rate (real num)

param

events per unit time

0, 1, 2, 3, ...

$X \sim N(\mu, \sigma)$

$X \sim Bin(n, p)$

discrete

real integer

8 / 8

poisson r.v

$X \sim \text{Poission}(\lambda)$ empirically

↳ website visit rate : 20 vis/hr

↳ 0, 1, 2, ...

limits in calc

PMF: $P(X=\boxed{k}) = \dfrac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, 2, \ldots$

100

CDF: $P(X \le k) = e^{-\lambda} \underbrace{\sum_{i=0}^{k} \frac{\lambda^i}{i!}}$ $= P(X=0) + P(X=1) + \cdots P(X=k)$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:30
9 / 9

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Examples_that_violate_the_Poisson_assumptions
Update

The Free Encyclopedia

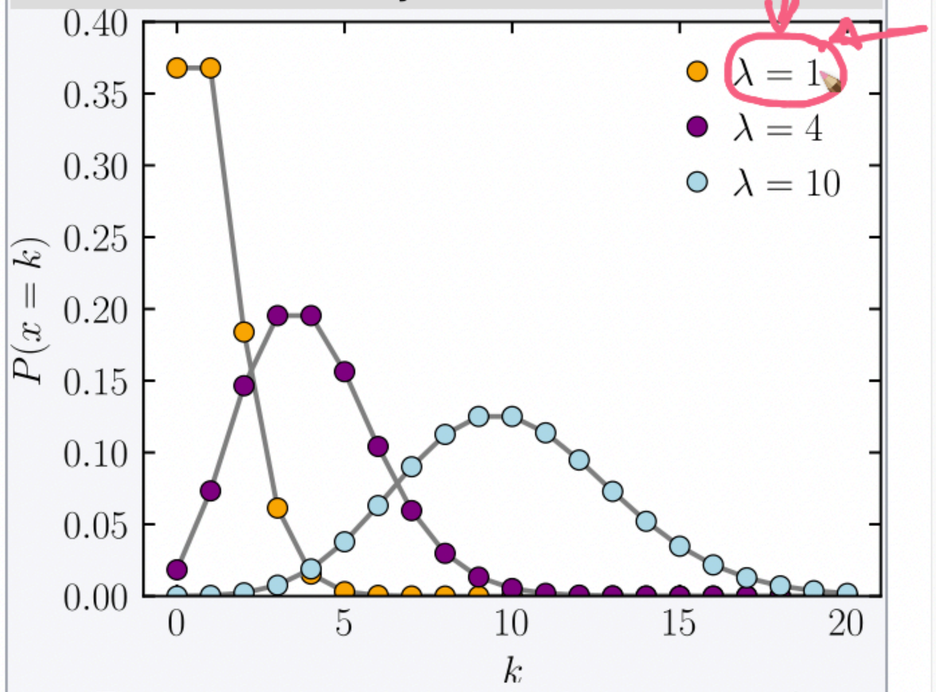
From Wikipedia, the free encyclopedia

In probability theory and statistics, the **Poisson distribution** is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.[1] It is named after French mathematician Siméon Denis Poisson (/ˈpwɑːsɒn/; French pronunciation: [pwasɔ̃]). The Poisson distribution can also be used for the number of events in other specified interval types such as distance, area or volume.

For instance, a call center receives an average of 180 calls per hour, 24 hours a day. The calls are independent; receiving one does not change the probability of when the next one will arrive. The number of calls received during any minute has a Poisson probability distribution: the most likely numbers are 2 and 3 but 1 and 4 are also likely and there is a small probability of it being as low as zero and a very small probability it could be 10. Another example is the number of decay events that occur from a radioactive source during a defined observation period.

Poisson Distribution

Probability mass function



The horizontal axis is the index *k*, the number of occurrences. λ is the expected rate of occurrences. The vertical axis is the probability of *k* occurrences given λ. The function is defined only at integer values of *k*; the connecting lines are only guides for the eye.

Cumulative distribution function

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe

en.wikipedia.org/wiki/Poisson_distribution#Examples_that_violate_the_Poisson_assumptions

The Free Encyclopedia

From Wikipedia, the free encyclopedia

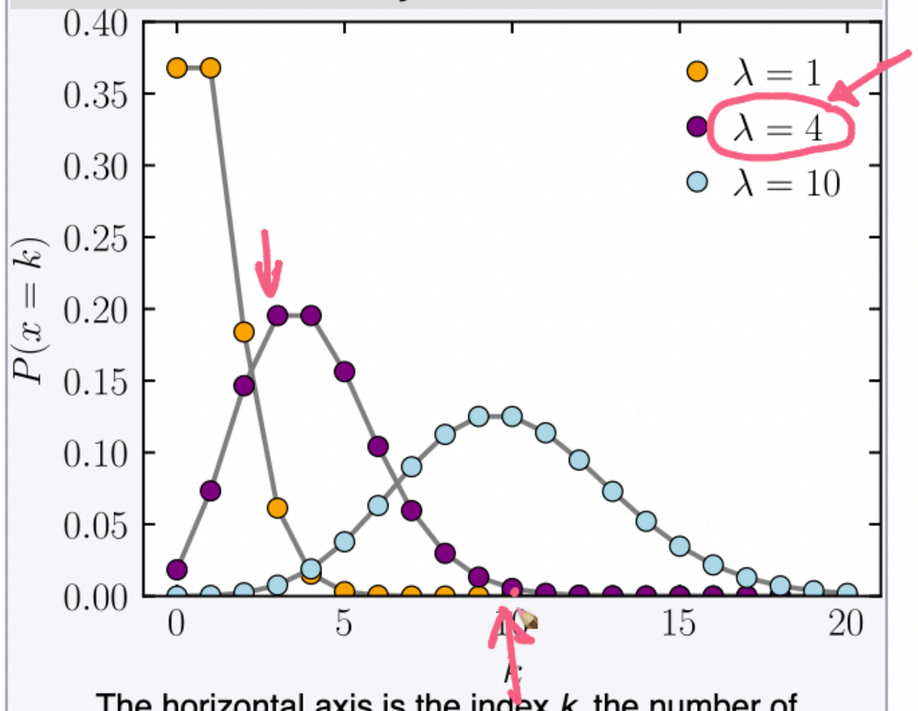
In probability theory and statistics, the **Poisson distribution** is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.[1] It is named after French mathematician Siméon Denis Poisson (/ˈpwɑːsɒn/; French pronunciation: [pwasɔ̃]). The Poisson distribution can also be used for the number of events in other specified interval types such as distance, area or volume.

For instance, a call center receives an average of 180 calls per hour, 24 hours a day. The calls are independent; receiving one does not change the probability of when the next one will arrive. The number of calls received during any minute has a Poisson probability distribution: the most likely numbers are 2 and 3 but 1 and 4 are also likely and there is a small probability of it being as low as zero and a very small probability it could be 10. Another example is the number of decay events that occur from a radioactive source during a defined observation period.

λ: 4vis/hr

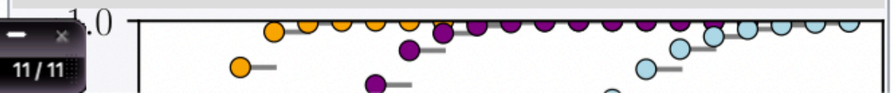
Poisson Distribution

Probability mass function



The horizontal axis is the index *k*, the number of occurrences. λ is the expected rate of occurrences. The vertical axis is the probability of *k* occurrences given λ. The function is defined only at integer values of *k*; the connecting lines are only guides for the eye.

Cumulative distribution function



Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Examples_that_violate_the_Poisson_assumptions
Update

The Free Encyclopedia

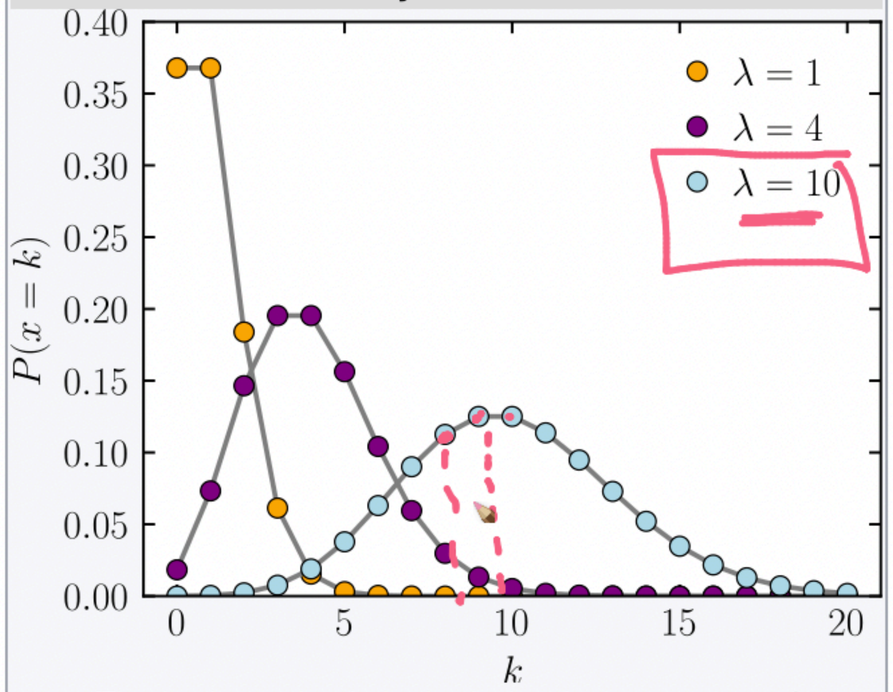
From Wikipedia, the free encyclopedia

In probability theory and statistics, the **Poisson distribution** is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.[1] It is named after French mathematician Siméon Denis Poisson (/ˈpwɑːsɒn/; French pronunciation: [pwasɔ̃]). The Poisson distribution can also be used for the number of events in other specified interval types such as distance, area or volume.

For instance, a call center receives an average of 180 calls per hour, 24 hours a day. The calls are independent; receiving one does not change the probability of when the next one will arrive. The number of calls received during any minute has a Poisson probability distribution: the most likely numbers are 2 and 3 but 1 and 4 are also likely and there is a small probability of it being as low as zero and a very small probability it could be 10. Another example is the number of decay events that occur from a radioactive source during a defined observation period.

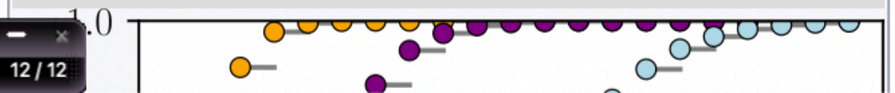
Poisson Distribution

Probability mass function



The horizontal axis is the index k, the number of occurrences. λ is the expected rate of occurrences. The vertical axis is the probability of k occurrences given λ. The function is defined only at integer values of k; the connecting lines are only guides for the eye.

Cumulative distribution function



12 / 12

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Examples_that_violate_the_Poisson_assumptions
Update

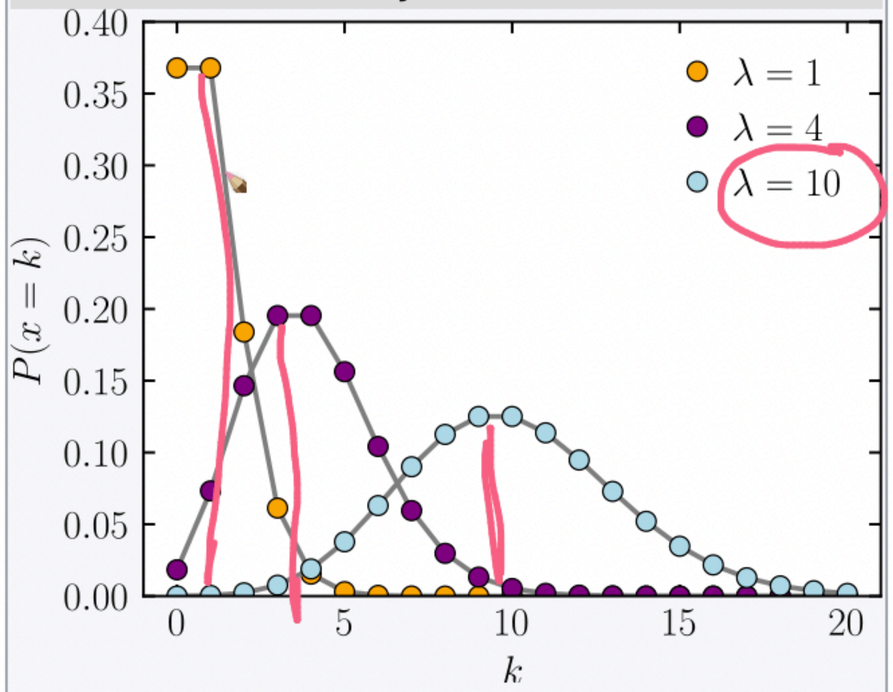
From Wikipedia, the free encyclopedia

In probability theory and statistics, the **Poisson distribution** is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.[1] It is named after French mathematician Siméon Denis Poisson (/ˈpwɑːsɒn/; French pronunciation: [pwasɔ̃]). The Poisson distribution can also be used for the number of events in other specified interval types such as distance, area or volume.

For instance, a call center receives an average of 180 calls per hour, 24 hours a day. The calls are independent; receiving one does not change the probability of when the next one will arrive. The number of calls received during any minute has a Poisson probability distribution: the most likely numbers are 2 and 3 but 1 and 4 are also likely and there is a small probability of it being as low as zero and a very small probability it could be 10. Another example is the number of decay events that occur from a radioactive source during a defined observation period.

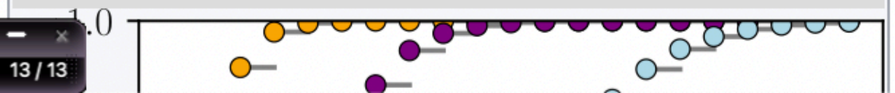
Poisson Distribution

Probability mass function



The horizontal axis is the index k, the number of occurrences. λ is the expected rate of occurrences. The vertical axis is the probability of k occurrences given λ. The function is defined only at integer values of k; the connecting lines are only guides for the eye.

Cumulative distribution function



Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Examples_that_violate_the_Poisson_assumptions

$\lambda = 2.5$

90 min

María Dolores Ugarte and colleagues report that the average number of goals in a World Cup soccer match is approximately 2.5 and the Poisson model is appropriate.[15] Because the average event rate is 2.5 goals per match, $\lambda = 2.5$.

$$P(k \text{ goals in a match}) = \frac{2.5^k e^{-2.5}}{k!}$$

$$P(k = 0 \text{ goals in a match}) = \frac{2.5^0 e^{-2.5}}{0!} = \frac{e^{-2.5}}{1} \approx 0.082$$

$$P(k = 1 \text{ goal in a match}) = \frac{2.5^1 e^{-2.5}}{1!} = \frac{2.5e^{-2.5}}{1} \approx 0.205$$

$$P(k = 2 \text{ goals in a match}) = \frac{2.5^2 e^{-2.5}}{2!} = \frac{6.25e^{-2.5}}{2} \approx 0.257$$

k	$P(k$ goals in a World Cup soccer match)
0	0.082
1	0.205
2	0.257
3	0.213
4	0.133
5	0.067
6	0.028
7	0.010

The probability for 0 to 7 goals in a match.

Once in an interval events: The special case of $\lambda = 1$ and $k = 0$

Suppose that ... size) hit the earth on average once every 100

14 / 16

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Examples_that_violate_the_Poisson_assumptions
Update

0 to 7 goals in a match.

$\lambda = 1$

unit time = 100 yrs

Once in an interval events: The special case of $\lambda = 1$ and $k = 0$

Suppose that astronomers estimate that large meteorites (above a certain size) hit the earth on average once every 100 years ($\lambda = 1$ event per 100 years), and that the number of meteorite hits follows a Poisson distribution. What is the probability of $k = 0$ meteorite hits in the next 100 years?

$$P(k = 0 \text{ meteorites hit in next 100 years}) = \frac{1^0 e^{-1}}{0!} = \frac{1}{e} \approx 0.37$$

Under these assumptions, the probability that no large meteorites hit the earth in the next 100 years is roughly 0.37. The remaining 1 – 0.37 = 0.63 is the probability of 1, 2, 3, or more large meteorite hits in the next 100 years. In an example above, an overflow flood occurred once every 100 years ($\lambda = 1$). The probability of no overflow floods in 100 years was roughly 0.37, by the same calculation.

In general, if an event occurs on average once per interval ($\lambda = 1$), and the events follow a Poisson distribution, then P(0 events in next interval) = 0.37. In addition, P(exactly one event in next interval) = 0.37, as shown in the table for overflow floods.

Examples that violate the Poisson assumptions

The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not con[illegible]es) and the arrivals of individual students are

15 / 17

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:38

groups

Examples that violate the Poisson assumptions

The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not constant (low rate during class time, high rate between class times) and the arrivals of individual students are not independent (students tend to come in groups).

The number of magnitude 5 earthquakes per year in a country may not follow a Poisson distribution if one large earthquake increases the probability of aftershocks of similar magnitude.

Examples in which at least one event is guaranteed are not Poisson distributed; but may be modeled using a zero-truncated Poisson distribution.

Count distributions in which the number of intervals with zero events is higher than predicted by a Poisson model may be modeled using a zero-inflated model.

Properties

Descriptive statistics

- The expected value and variance of a Poisson-distributed random variable are both equal to λ.
- The coefficient of variation is $\lambda^{-1/2}$, while the index of dispersion is 1.[6]:163
- The mean absolute deviation about the mean is[6]:163

$$\mathrm{E}[|X - \lambda|] = \frac{2\lambda^{\lfloor\lambda\rfloor+1}e^{-\lambda}}{}$$

16 / 18

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:39
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Occurrence_and_applications
Update

challenged and removed.

Find sources: "Poisson distribution" – news · newspapers · books · scholar · JSTOR *(December 2019) (Learn how and when to remove this template message)*

>1000calls/min → v.v. low

>100calls/min

Applications of the Poisson distribution can be found in many fields including:[45]

- count data in general
- Telecommunication example: telephone calls arriving in a system. ← cell-tower
- Astronomy example: photons arriving at a telescope.
- Chemistry example: the molar mass distribution of a living polymerization.[46]
- Biology example: the number of mutations on a strand of DNA per unit length.
- Management example: customers arriving at a counter or call centre.
- Finance and insurance example: number of losses or claims occurring in a given period of time.
- Earthquake seismology example: an asymptotic Poisson model of seismic risk for large earthquakes.[47]
- Radioactivity example: number of decays in a given time interval in a radioactive sample.
- Optics example: the number of photons emitted in a single laser pulse. This is a major vulnerability to most Quantum key distribution protocols known as Photon Number Splitting (PNS).

The Poisson distribution arises in connection with Poisson processes. It applies to various phenomena of discrete properties (that is, those that may happen 0, 1, 2, 3, … times during a given period of time or in a given area) whenever the probability of the phenomenon happening is constant in time or space. Examples of events that may be modelled as a Poisson distribution include:

- The numb … the Prussian cavalry. This example was used

17 / 19

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:40
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Occurrence_and_applications
Update

challenged and removed.

Find sources: "Poisson distribution" – news · newspapers · books · scholar · JSTOR *(December 2019)* (*Learn how and when to remove this template message*)

Applications of the Poisson distribution can be found in many fields including:[45]

- count data in general
- Telecommunication example: telephone calls arriving in a system.
- Astronomy example: photons arriving at a telescope.
- Chemistry example: the molar mass distribution of a living polymerization.[46]
- Biology example: the number of mutations on a strand of DNA per unit length.
- Management example: customers arriving at a counter or call centre.
- Finance and insurance example: number of losses or claims occurring in a given period of time.
- Earthquake seismology example: an asymptotic Poisson model of seismic risk for large earthquakes.[47]
- Radioactivity example: number of decays in a given time interval in a radioactive sample. 1-day
- Optics example: the number of photons emitted in a single laser pulse. This is a major vulnerability to most Quantum key distribution protocols known as Photon Number Splitting (PNS).

The Poisson distribution arises in connection with Poisson processes. It applies to various phenomena of discrete properties (that is, those that may happen 0, 1, 2, 3, … times during a given period of time or in a given area) whenever the probability of the phenomenon happening is constant in time or space. Examples of events that may be modelled as a Poisson distribution include:

- The numb … the Prussian cavalry. This example was used

18 / 20

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:41

Exponential

19 / 21

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:42
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Occurrence_and_applications
Update

What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Print/export

Download as PDF
Printable version

In other projects

Wikimedia Commons

Languages

العربية
বাংলা
Español
Bahasa Indonesia
Bahasa Melayu
Português
Русский
اردو

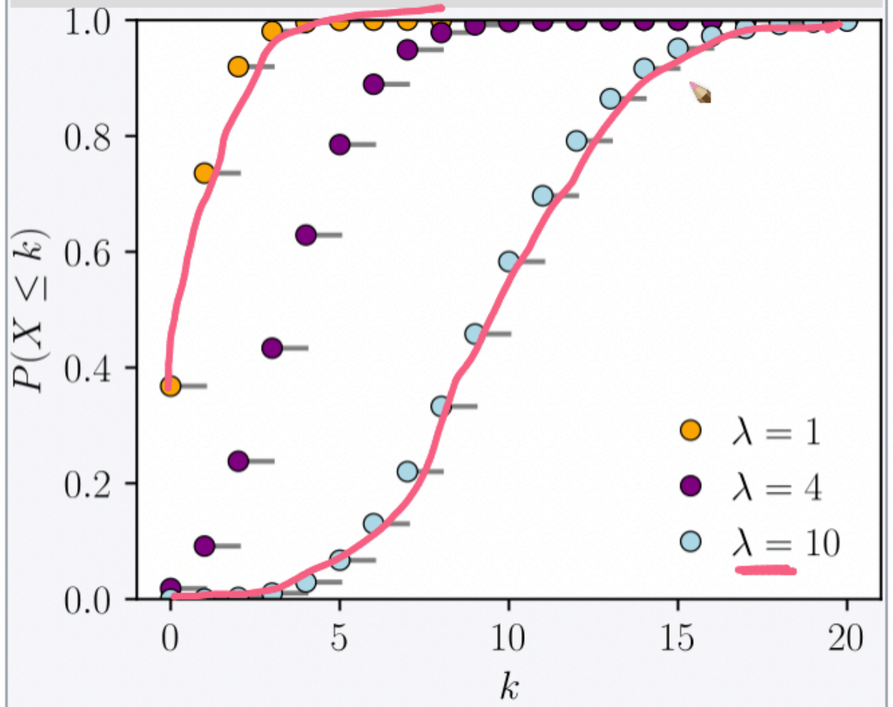
probability distribution: the most likely numbers are 2 and 3 but 1 and 4 are also likely and there is a small probability of it being as low as zero and a very small probability it could be 10. Another example is the number of decay events that occur from a radioactive source during a defined observation period.

Contents [hide]

1 History
2 Definitions
2.1 Probability Mass Function
2.2 Example
2.3 Assumptions and validity
2.3.1 Examples of probability for Poisson distributions
2.3.2 Once in an interval events: The special case of $\lambda = 1$ and $k = 0$
2.4 Examples that violate the Poisson assumptions
3 Properties
3.1 Descriptive statistics
3.2 Median
3.3 Higher moments
3.4 Sums of Poisson-distributed random variables
3.5 Othe

The horizontal axis is the index k, the number of occurrences. λ is the expected rate of occurrences. The vertical axis is the probability of k occurrences given λ. The function is defined only at integer values of k; the connecting lines are only guides for the eye.

Cumulative distribution function



The horizontal axis is the index k, the number of occurrences. The CDF is discontinuous at the integers of k and flat everywhere else because a variable that is Poisson distributed takes on only integer values.

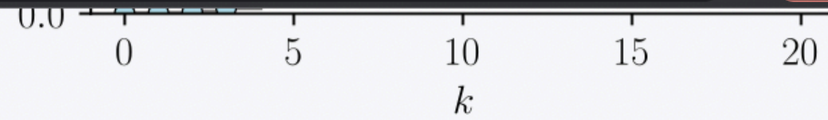
Notation	$\mathrm{Pois}(\lambda)$
meters	$\lambda \in (0, \infty)$ (rate)

20 / 22

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:42
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Occurrence_and_applications
Update

العربية
বাংলা
Español
Bahasa Indonesia
Bahasa Melayu
Português
Русский
اردو
中文
40 more
Edit links

2.4 Examples that violate the Poisson assumptions
3 Properties
3.1 Descriptive statistics
3.2 Median
3.3 Higher moments
3.4 Sums of Poisson-distributed random variables
3.5 Other properties
3.6 Poisson races
4 Related distributions
4.1 General
4.2 Poisson approximation
4.3 Bivariate Poisson distribution
4.4 Free Poisson distribution
4.4.1 Some transforms of this law
5 Statistical inference
5.1 Parameter estimation
5.2 Confidence interval
5.3 Bayesian inference
5.4 Simultaneous estimation of multiple Poisson means
6 Occurrence and applications
6.1 Law of rare events
6.2 Pois



The horizontal axis is the index *k*, the number of occurrences. The CDF is discontinuous at the integers of *k* and flat everywhere else because a variable that is Poisson distributed takes on only integer values.

Notation	$\mathrm{Pois}(\lambda)$
Parameters	$\lambda \in (0, \infty)$ (rate)
Support	$k \in \mathbb{N}_0$ (Natural numbers starting from 0)
PMF	$\frac{\lambda^k e^{-\lambda}}{k!}$
CDF	$\frac{\Gamma(\lfloor k+1 \rfloor, \lambda)}{\lfloor k \rfloor !}$, or $\boxed{e^{-\lambda} \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}}$, or $Q(\lfloor k+1 \rfloor, \lambda)$ (for $k \geq 0$, where $\Gamma(x, y)$ is the upper incomplete gamma function, $\lfloor k \rfloor$ is the floor function, and Q is the regularized gamma function)
Mean	λ
Median	$\approx \lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$
Mode	$\lceil \lambda \rceil - 1, \lfloor \lambda \rfloor$
Variance	λ
vness	$\lambda^{-1/2}$

21 / 23

Exponential dist

Const rate(λ) of events

events are indep

no simultaneous events

Poisson process

↳ satisfies the 3 conditions that underlie poisson dist

time interval between 2 events follows expo-dist

(e.g) goals in a WC-soccer match (wiki) → poisson process ($\lambda = 2.5$)

goals per match ~ Poisson ($\lambda = 2.5$)

time-interval between 2 goals ~ Expo(λ) → rate

(or) ~~Discrete~~ Continous

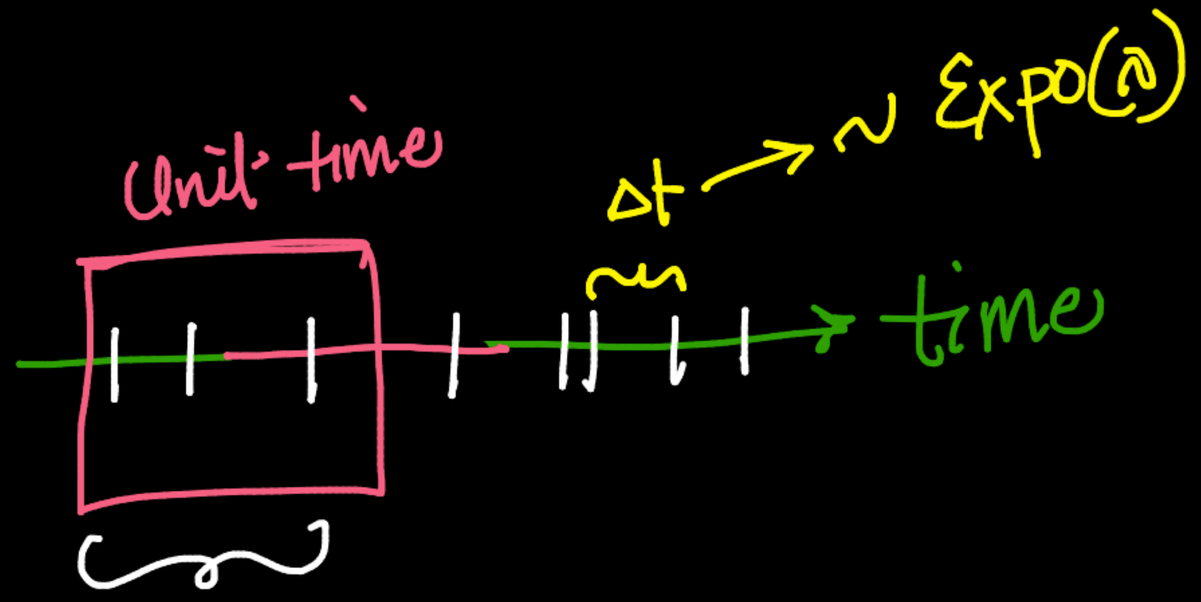
$X \sim Expo(\lambda)$

Poisson-process (3-conditions)

→ # of events per unit time → Poisson(λ)

→ Δt b/w 2 events → Expo(λ)

Poisson-process:



events

↓

Poisson(λ)

1: event

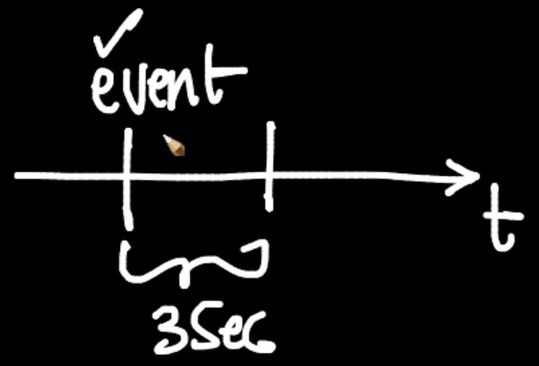
Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:04

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

Δt

PDF: $P(X = \underline{x}) = \lambda e^{-\lambda x}$

3sec



CDF: $P(X \leq x) = 1 - e^{-x\lambda}$

Poisson-process with $\lambda = 2$

26 / 34

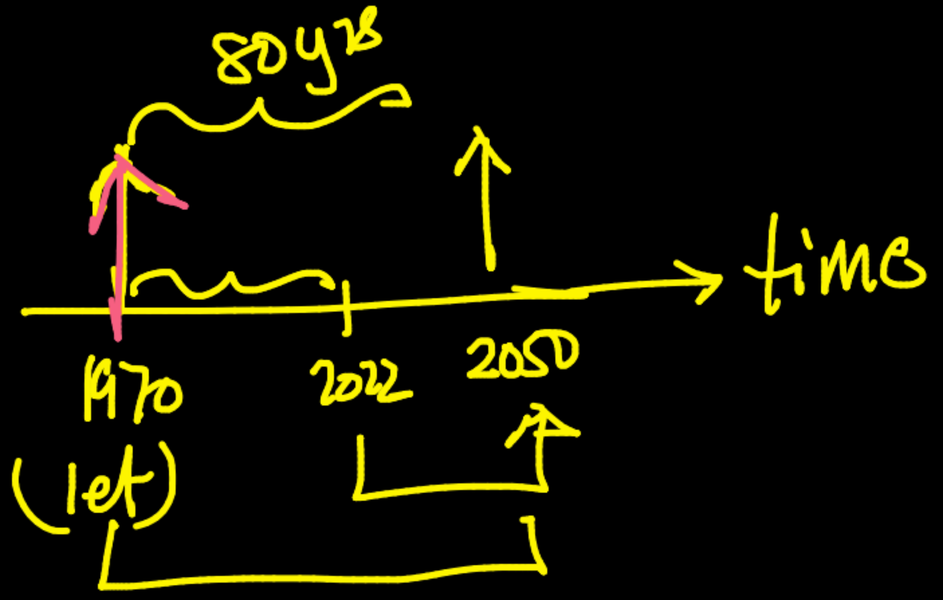
Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:08

$$P(X = \boxed{x}) = \lambda e^{-\lambda x}$$

$x \to \Delta t$

$$P(X = 80) = 1 \cdot e^{-1 \times 80} = \boxed{}$$

Poisson process : meteorites hitting (wiki) : $\boxed{1}$ per 100 yrs $= \lambda$



Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:55
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Exponential_distribution
Update

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Print/export
Download as PDF

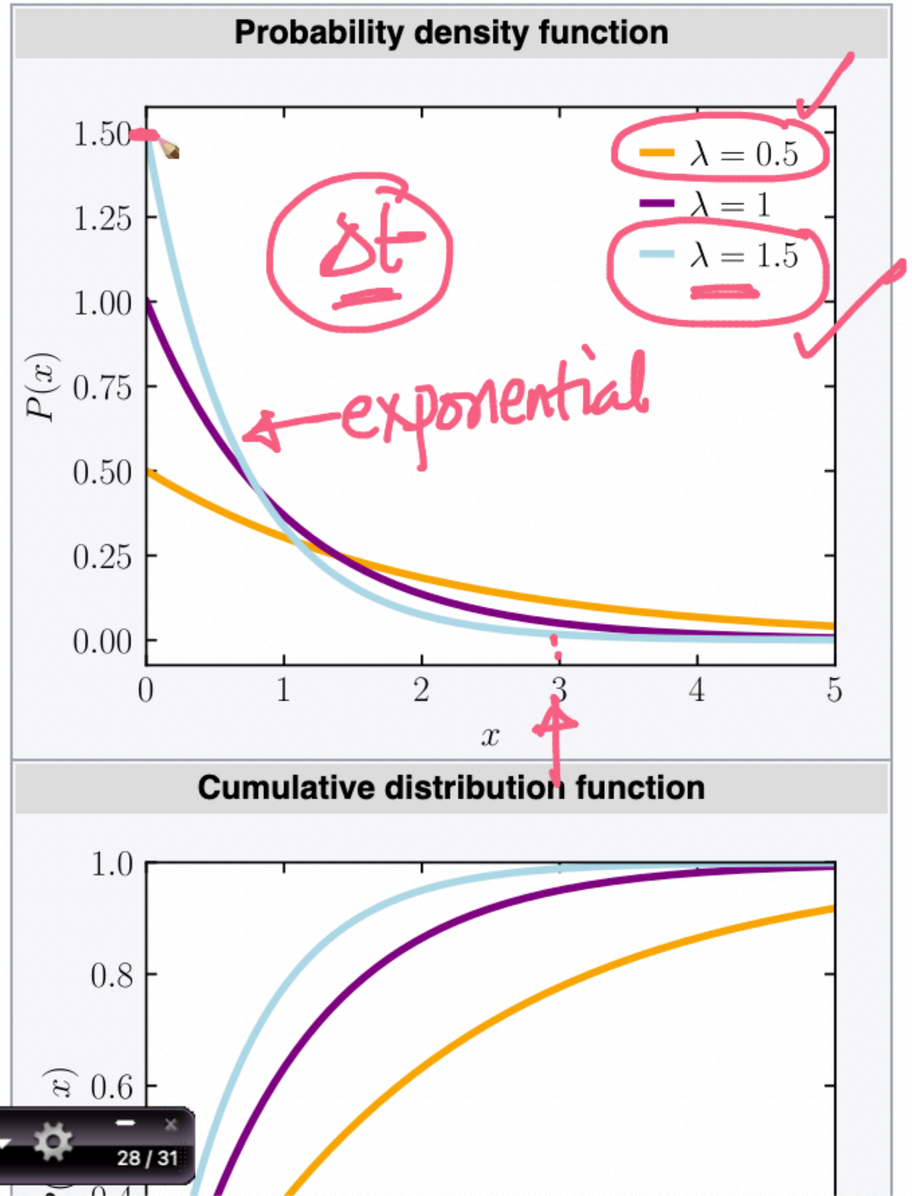
In probability theory and statistics, the **exponential distribution** is the probability distribution of the time between events in a Poisson point process, i.e., a process in which events occur continuously and independently at a constant average rate. It is a particular case of the gamma distribution. It is the continuous analogue of the geometric distribution, and it has the key property of being memoryless. In addition to being used for the analysis of Poisson point processes it is found in various other contexts.

The exponential distribution is not the same as the class of exponential families of distributions, which is a large class of probability distributions that includes the exponential distribution as one of its members, but also includes the normal distribution, binomial distribution, gamma distribution, Poisson, and many others.

Contents [hide]

1 Definitions
1.1 Probability density function
1.2 Cumulative distribution function
1.3 Alternative parametrization
2 Properties

Exponential



Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:57
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Exponential_distribution#Occurrence_and_applications
Update

Occurrence of events

The exponential distribution occurs naturally when describing the lengths of the inter-arrival times in a homogeneous Poisson process.

The exponential distribution may be viewed as a continuous counterpart of the geometric distribution, which describes the number of Bernoulli trials necessary for a *discrete* process to change state. In contrast, the exponential distribution describes the time for a continuous process to change state.

In real-world scenarios, the assumption of a constant rate (or probability per unit time) is rarely satisfied. For example, the rate of incoming phone calls differs according to the time of day. But if we focus on a time interval during which the rate is roughly constant, such as from 2 to 4 p.m. during work days, the exponential distribution can be used as a good approximate model for the time until the next phone call arrives. Similar caveats apply to the following examples which yield approximately exponentially distributed variables:

- The time until a radioactive particle decays, or the time between clicks of a Geiger counter
- The time it takes before your next telephone call
- The time until default (on payment to company debt holders) in reduced-form credit risk modeling

Exponential variables can also be used to model situations where certain events occur with a constant probability per unit length, such as the distance between mutations on a DNA strand, or between roadkills on a given road.

In queuing theory, the service times of agents in a system (e.g. how long it takes for a bank teller etc. to serve a customer) are often modeled as exponentially distributed variables. (The arrival of customers for instance is also modeled by the Poisson distribution [illegible] length of a process that can be thought

29 / 32

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:01

#visitors per min to your website b/w 5PM-9PM

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

1000's → Servers

Δt between 2 visitors $\sim \text{Expo}(\lambda)$

→ Auto scaling

2AM

30 / 33

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:03

3 conditions

Poisson process

→ # events per unit time ~ Poisson(λ)

→ Δt between events ~ Expo(λ)

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:03

Exponential: memoryless

32 / 34

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:10

Exponential: memoryless [property]

λ

web-traffic

Δt

$\hookrightarrow \text{Expo}(\lambda)$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:14
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Poisson_distribution#Occurrence_and_applications
mutations 1/3

article by adding citations to reliable sources. Unsourced m
challenged and removed.

Find sources: "Poisson distribution" – news · newspapers · books · scholar · JSTOR *(December 2019) (Learn how and when to remove this template message)*

Δt between mutation ~Expo(λ)

Applications of the Poisson distribution can be found in many fields including:[45]

- count data in general
- Telecommunication example: telephone calls arriving in a system.
- Astronomy example: photons arriving at a telescope.
- Chemistry example: the molar mass distribution of a living polymerization.[46]
- Biology example: the number of mutations on a strand of DNA per unit length.
- Management example: customers arriving at a counter or call centre.
- Finance and insurance example: number of losses or claims occurring in a given period of time.
- Earthquake seismology example: an asymptotic Poisson model of seismic risk for large earthquakes.[47]
- Radioactivity example: number of decays in a given time interval in a radioactive sample.
- Optics example: the number of photons emitted in a single laser pulse. This is a major vulnerability to most Quantum key distribution protocols known as Photon Number Splitting (PNS).

mutations in DNA (poisson-process (λ))

The Poisson distribution arises in connection with Poisson processes. It applies to various phenomena of discrete properties (that is, those that may happen 0, 1, 2, 3, … times during a given period of time or in a given area) whenever the probability of the ples of events that may be modelled as a Poisson distribution include:

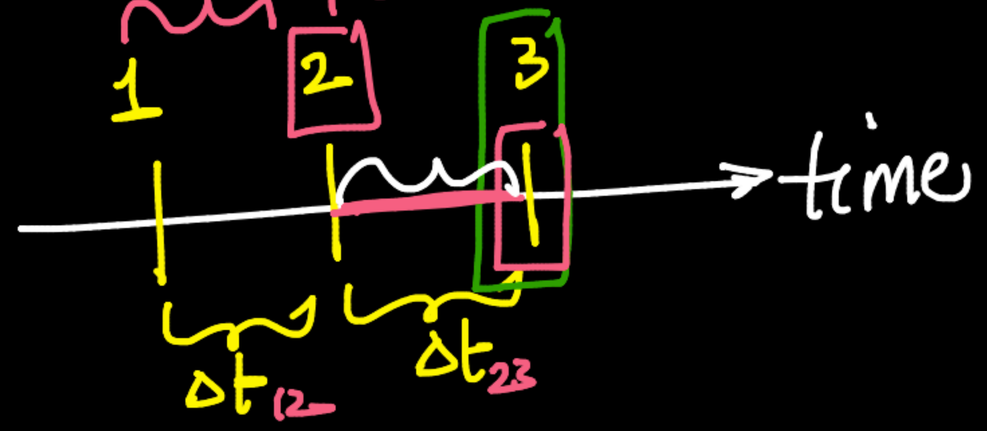
34 / 35

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:19

$\Delta t_{12} \overset{X}{>} \Delta t_{23}$

$\Delta t \sim \text{Expo}(\lambda = 2.5)$

memoryless



{ opposition / teams goals }

goals in a WC-soccer match: poisson process (wiki) approx... }

35 / 35

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:35

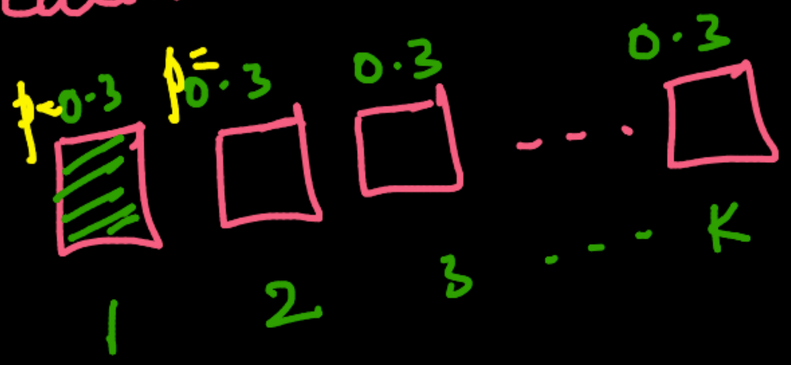
Geometric Distb

$X \sim$ number of failures before a success in a series of bernoulli trails

$p = 0.3$

e.g.: job-applications

LinkedIn Jobs / Jobsportal →



36 / 42

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:31

Conditions :-

1. Sequence of indep. trails
2. success & failure events only
   ↳ bernouli trails
3. p = prob. of success is same for each trail

37 / 42

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:25

$X \sim \text{Geom}(p)$ → prob. of success in a trail

(e.g) $\boxed{X} \sim$ Number of failed job applications before you succeed

$\{ p = 0.3$ (let)

38 / 39

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:29

$p = 0.3$

$P(X=5) =$ prob that we fail 5 job applications before we succeed in the 6th one

F	F	F	F	F	S
1	2	3	4	5	6

$(1-p)^5 p^1$

$0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.7 \times 0.3$

$= P(F_1 \text{ and } F_2 \text{ and } F_3 \text{ and } F_4 \text{ and } F_5 \text{ and } S_6)$

$= P(F_1) * P(F_2) * P(F_3) * P(F_4) * P(F_5) * P(S_6)$

39 / 40

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:28

$$\left\{ \begin{array}{l} P(X \text{ and } Y) = P(X) * P(Y) \\ X \text{ is } \underline{\underline{\text{indep}}} \text{ of } Y \end{array} \right\} \leftarrow \text{Conditional prob}$$

40 / 40

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:35

PMF

$$P(X = \boxed{k}) = (1-p)^k \cdot p^1$$

$X \sim Geom(p)$

prob. of failing k-times & succeeding in the $k+1^{th}$ time

CDF: — $$P(X \le k) = \sum_{i=1}^{k} (1-p)^i p^1$$

41 / 42

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:34

examples:

Sales - success

$(p = 0.1)$

1. # tosses to make before you get a head (HT)

2. T-shirt fitting (XL)

3. RGB →

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:36

22:40

43 / 43

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedi | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Geometric_distribution#Probability_outcomes_examples
Update

for $k = 0, 1, 2, 3, \ldots$.

E1) A doctor is seeking an antidepressant for a newly diagnosed patient. Suppose that, of the available anti-depressant drugs, the probability that any particular drug will be effective for a particular patient is $p = 0.6$. What is the probability that the first drug found to be effective for this patient is the first drug tried, the second drug tried, and so on? What is the expected number of drugs that will be tried to find one that is effective?

The probability that the first drug works. There are zero failures before the first success. $Y = 0$ failures. The probability P(zero failures before first success) is simply the probability that the first drug works.

$$\Pr(Y = 0) = q^0\, p \;= 0.4^0 \times 0.6 = 1 \times 0.6 = 0.6.$$

The probability that the first drug fails, but the second drug works. There is one failure before the first success. Y= 1 failure. The probability for this sequence of events is P(first drug fails) $\times$ p(second drug succeeds), which is given by

$$\Pr(Y = 1) = q^1\, p \;= 0.4^1 \times 0.6 = 0.4 \times 0.6 = 0.24.$$

The probability that the first drug fails, the second drug fails, but the third drug works. There are two failures before the first success. $Y = 2$ failures. The probability for this sequence of events is P(first drug fails) $\times$ p(second drug fails) $\times$ P(third drug is success)

$$\Pr(Y = 2) = q^2\, p, = 0.4^2 \times 0.6 = 0.096.$$

E2) A newlywed couple plans to have children and will continue until the first girl. What is the probability that there are zero boys before the first girl, one boy before the first girl, two boys before the first girl, and so on?

The probability of ha boy (failure) is $q = 1 - p = 0.5$.

44 / 44

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Geometric_distribution#Probability_outcomes_examples
Update

The probability that the first drug fails, the second drug fails, but the third drug works. There are two failures before the first success. $Y = 2$ failures. The probability for this sequence of events is P(first drug fails) $\times$ p(second drug fails) $\times$ P(third drug is success)

$$\Pr(Y = 2) = q^2\, p, = 0.4^2 \times 0.6 = 0.096.$$

E2) A newlywed couple plans to have children and will continue until the first girl. What is the probability that there are zero boys before the first girl, one boy before the first girl, two boys before the first girl, and so on?

The probability of having a girl (success) is p= 0.5 and the probability of having a boy (failure) is $q = 1 - p = 0.5$.

The probability of no boys before the first girl is

$$\Pr(Y = 0) = q^0\, p \ = 0.5^0 \times 0.5 = 1 \times 0.5 = 0.5.$$

The probability of one boy before the first girl is

$$\Pr(Y = 1) = q^1\, p \ = 0.5^1 \times 0.5 = 0.5 \times 0.5 = 0.25.$$

The probability of two boys before the first girl is

$$\Pr(Y = 2) = q^2\, p \ = 0.5^2 \times 0.5 = 0.125.$$

and so on.

Properties [edit]

Moments and cumulants [edit]

The expected value for the number of independent trials to get the first success, and the variance of a geometrically

45 / 45

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Poisson distribution - Wikipedia | Exponential distribution - Wiki | Geometric distribution - Wikip | Log-normal distribution - Wiki | Bootstrapping (statistics) - Wi | Central limit theorem - Wikipe
en.wikipedia.org/wiki/Geometric_distribution#Probability_outcomes_examples
Update

The probability that the first drug fails, the second drug fails, but the third drug works. There are two failures before the first success. $Y = 2$ failures. The probability for this sequence of events is P(first drug fails) $\times$ p(second drug fails) $\times$ P(third drug is success)

$$\Pr(Y = 2) = q^2\, p, = 0.4^2 \times 0.6 = 0.096.$$

E2) A newlywed couple plans to have children and will continue until the first girl. What is the probability that there are zero boys before the first girl, one boy before the first girl, two boys before the first girl, and so on?

The probability of having a girl (success) is p= 0.5 and the probability of having a boy (failure) is $q = 1 - p = 0.5$.

The probability of no boys before the first girl is

$$\Pr(Y = 0) = q^0\, p \ = 0.5^0 \times 0.5 = 1 \times 0.5 = 0.5.$$

The probability of one boy before the first girl is

$$\Pr(Y = 1) = q^1\, p \ = 0.5^1 \times 0.5 = 0.5 \times 0.5 = 0.25.$$

The probability of two boys before the first girl is

$$\Pr(Y = 2) = q^2\, p \ = 0.5^2 \times 0.5 = 0.125.$$

and so on.

Properties [edit]

Moments and cumulants [edit]

The expected value for the number of independent trials to get the first success, and the variance of a geometrically

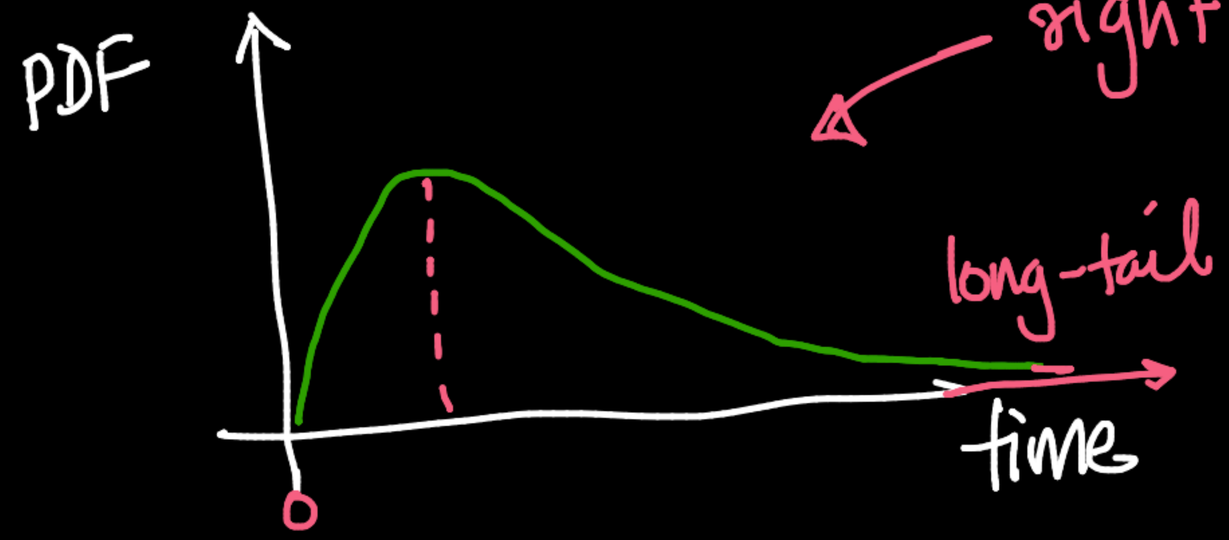
46 / 46

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:44

Continous $X \sim \text{Log-normal}(\mu, \sigma)$

$\boxed{\text{Log-normal}}$

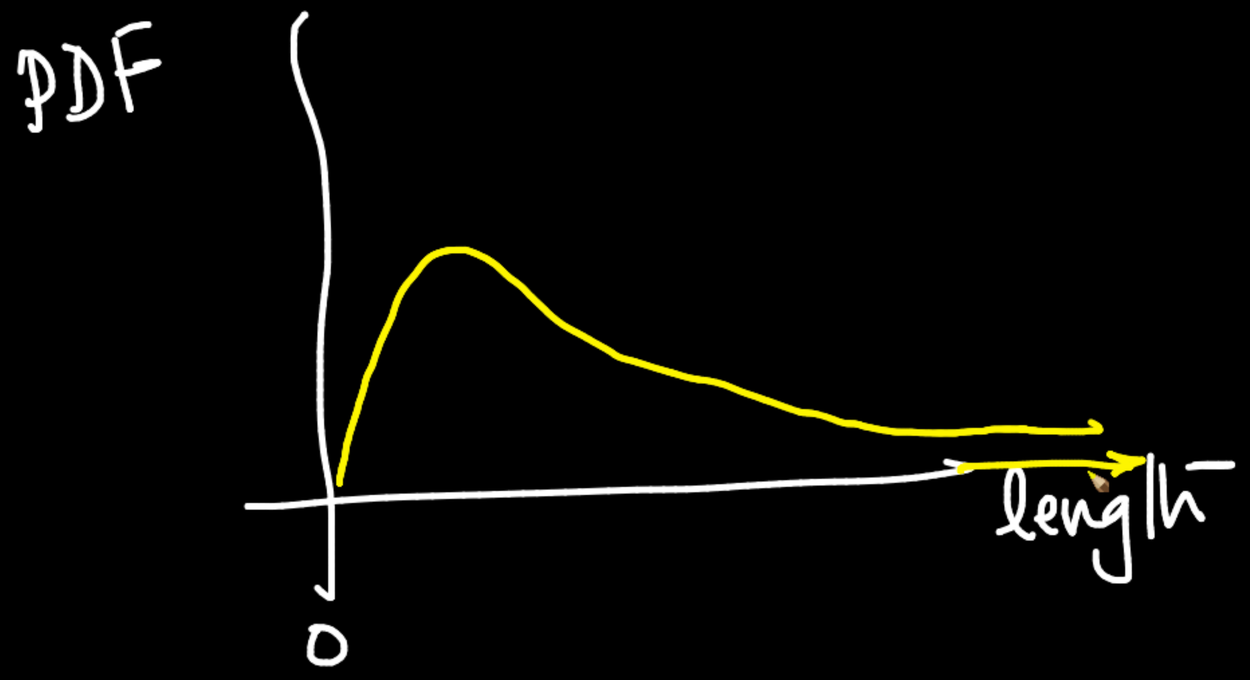
(e.g) delivery-time of Zomato / Amazon



47 / 47

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:45

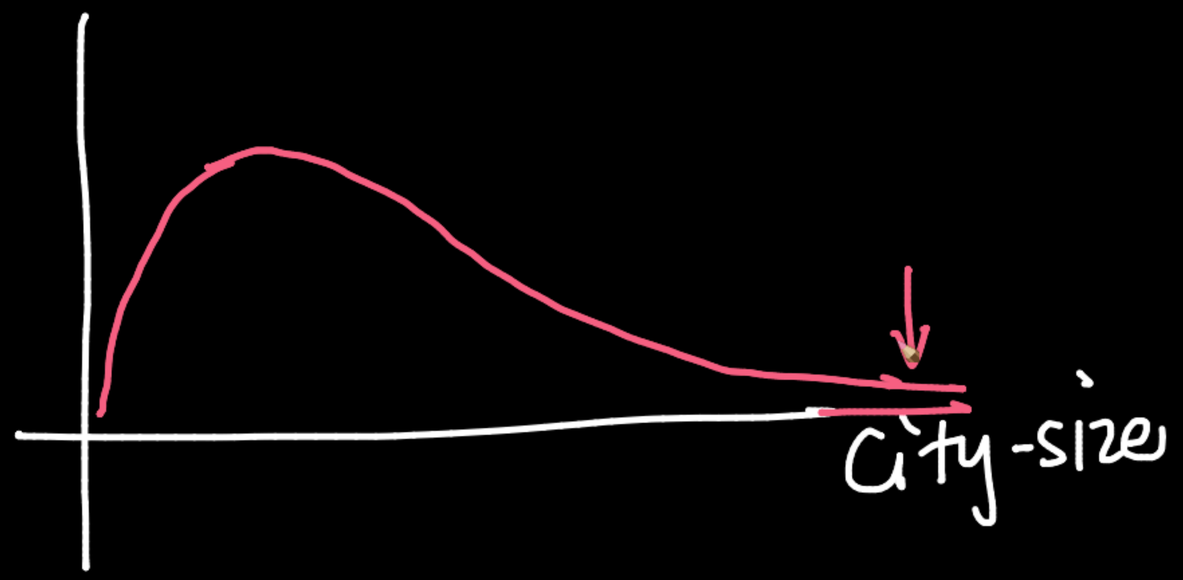
(e.g) comments-length on Youtube



48 / 48

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:46

(e.g) Size of cities



49 / 49

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:56

Math:

$X \sim \text{log-normal}(\mu, \sigma)$ ≥ 0

$\Updownarrow$

Def

If $\log_e(X) \sim \text{Normal}(\mu, \sigma)$

Condition

Obs: (delivery-times)

$\boxed{x_1, x_2, x_3, \ldots x_{1000}}$

$\log(x_1), \log(x_2) \ldots \log(x_{1000})$

$\downarrow x_1' \quad \downarrow x_2' \quad \ldots \quad \downarrow x_{1000}'$

If $x_i' \sim \text{Normal}(\mu, \sigma)$

then

x_is $\sim$ log-normal (μ, σ)

Statistical tests

Normal-distb → occurs a nature

log-Normal → also occurs a nature

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:50

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:56

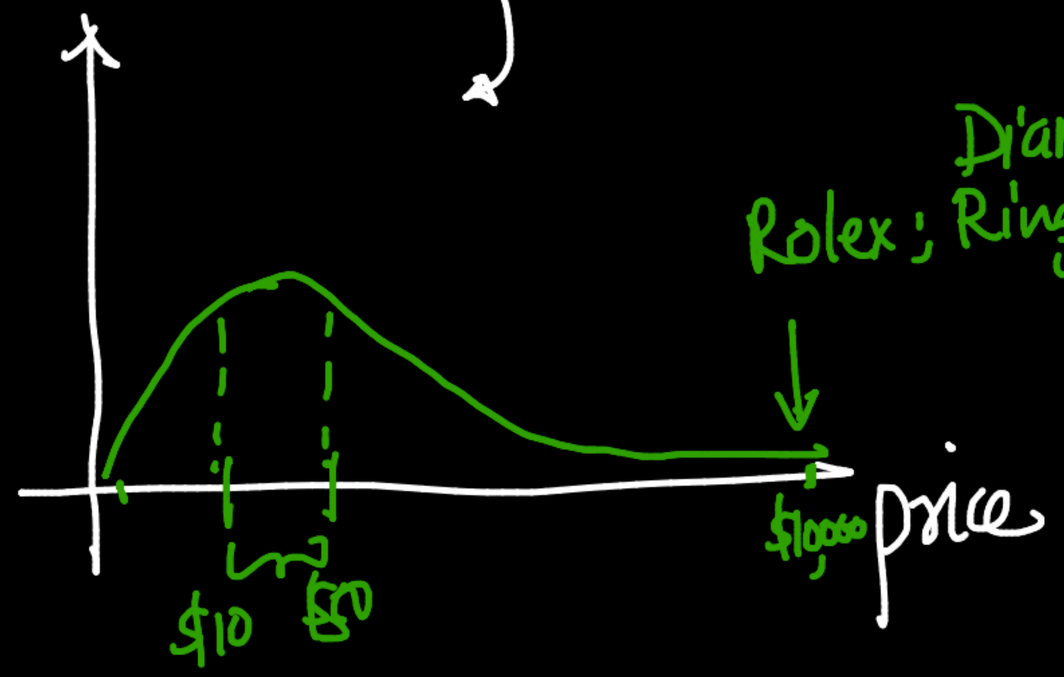
{ pareto-dist

prices of products on e-commerce

- intuition (approx) → guess
- certain accurate

log(prices)

{ statistical tests & plots



Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:00

(Q) Number of patients with side effects from a medicine from a group of 100 patients → p

X geometric → k-failures before a success

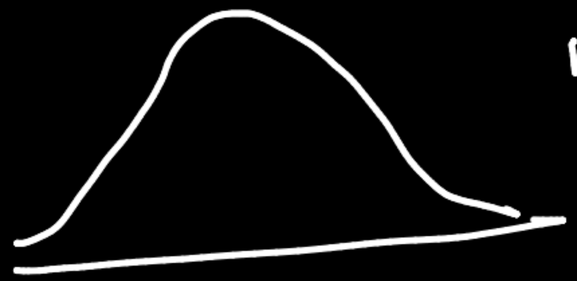
Binomial → ✓

53 / 53

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:02

Q

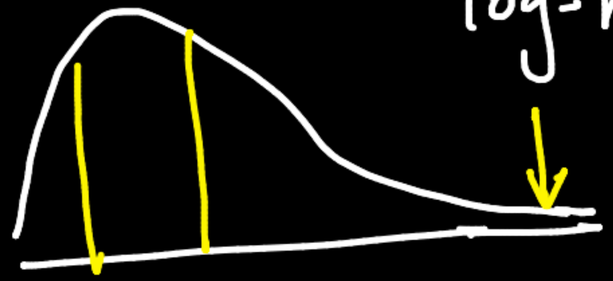
Incomes of working adults



normal: up

certain:
log(incomes)
& stat tests

guess:



log-normal: down

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:05

Q

QC @ an industry:

#items to inspect before detecting a defective product

success

Seq. of bernoulli trails

Geometric

$p = 0.01$

✓ ✓ ✗

55 / 55

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:09

Ex/ Inspection process: → const-rate $\boxed{(\lambda)}$ of finding defective products

{ Poisson-process: → Const-rate (λ) ✓

indep ✓

no sim-events ✓

{ Q X = # defective pieces found in unit time → $\boxed{\text{Poisson}(\lambda)}$ ✓ ✓

56 / 56

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:16

data → $\bar{x}$: mean

s: std-dev

obs → $x_1, x_2, x_3, \ldots x_{100}$ ①

X heights

Check/ Test

$\sim$ Normal$(\mu = 160, \sigma = 10)$ ② ③

Ideas:

① z-score: $\dfrac{x_i - \bar{x}}{s}$

Good but not correct → ② 68-95-99 →

$[\mu - \sigma, \mu + \sigma]$ 150 170

$[\mu - 2\sigma, \mu + 2\sigma]$

$[\mu - 3\sigma, \mu + 3\sigma]$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:15

$X = x_1, x_2 \ldots x_{100}$

$Y \sim \text{Normal}(\mu, \sigma)$

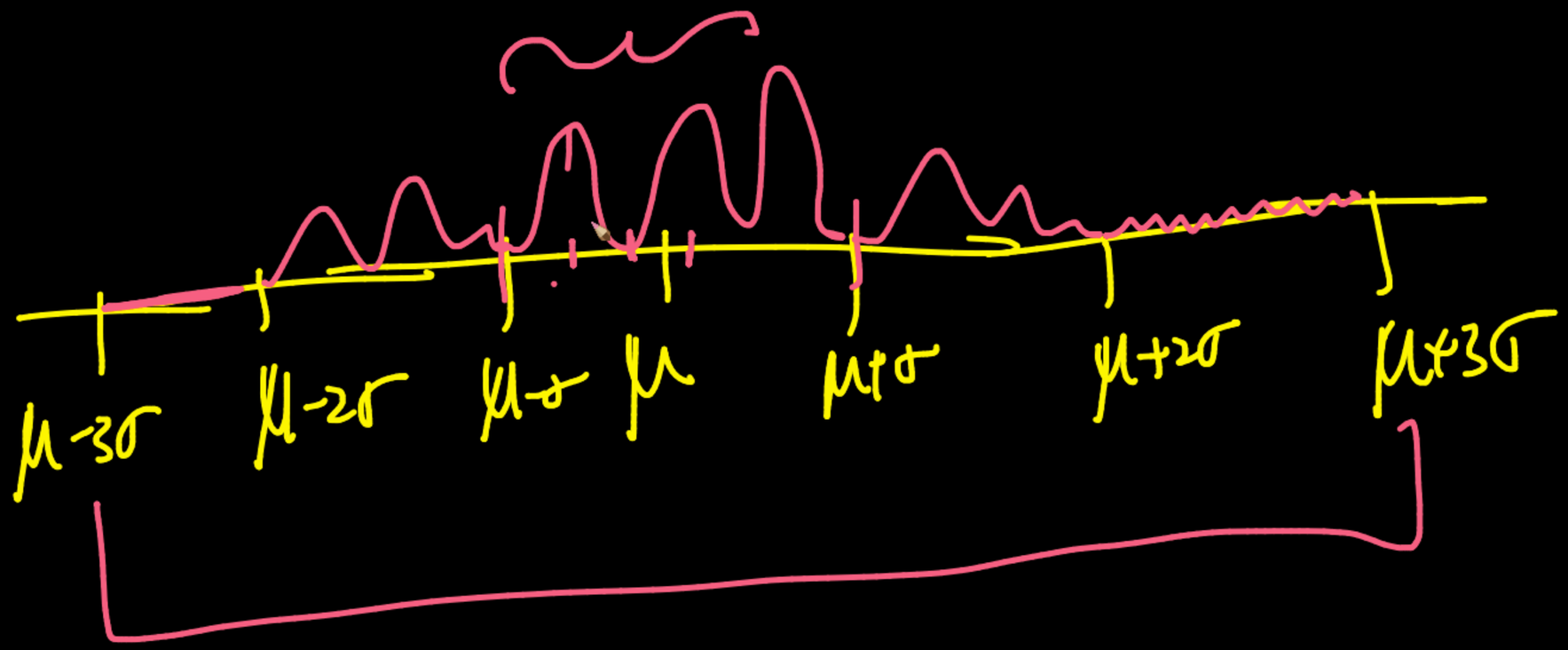
68-95-99 rule

✓ If $Y \sim$ Normal then 68-95-99 rule

✗ If 68-95-99 rule then Gaussian

58 / 58

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:16



59 / 59

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:20

$x_1 \quad x_2 \quad \cdots \quad x_{100}$

Normal ($\mu = 160$, $\sigma = 10$)

✓ 1

mean $\bar{x}$

sd-dev s

not solid test

$\bar{x} \sim 160$

$s \sim 10$

median = 160

2

median of 100-values

less impacted to outliers

soft percentile

159 let

~ 160

NOT definitive

60 / 61

$x_1, x_2 \ldots\ldots x_{100}$

$\text{Normal}(\mu = 160, \sigma = 10)$

box-plot →

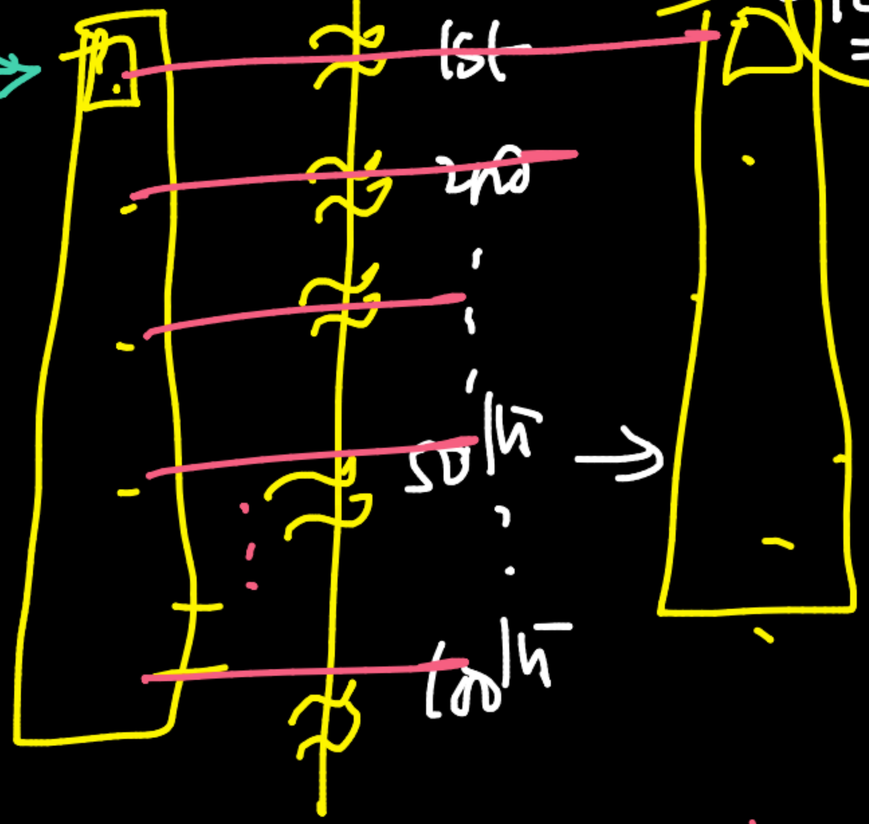
1st percentile →

2nd →

3rd →

50th →

100th →



100,000 pts

vs

Very confident

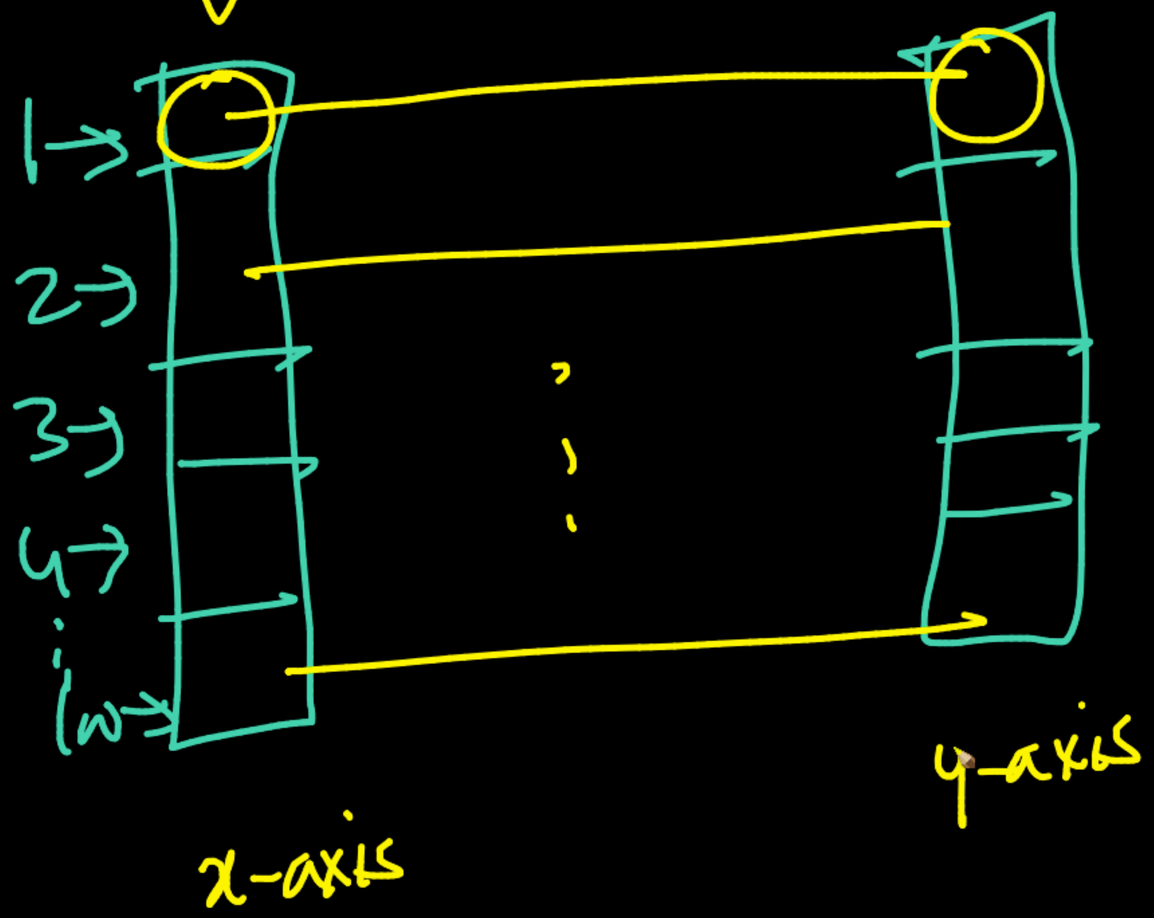
Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:26

(data) observed percentiles

Theoretical percentiles

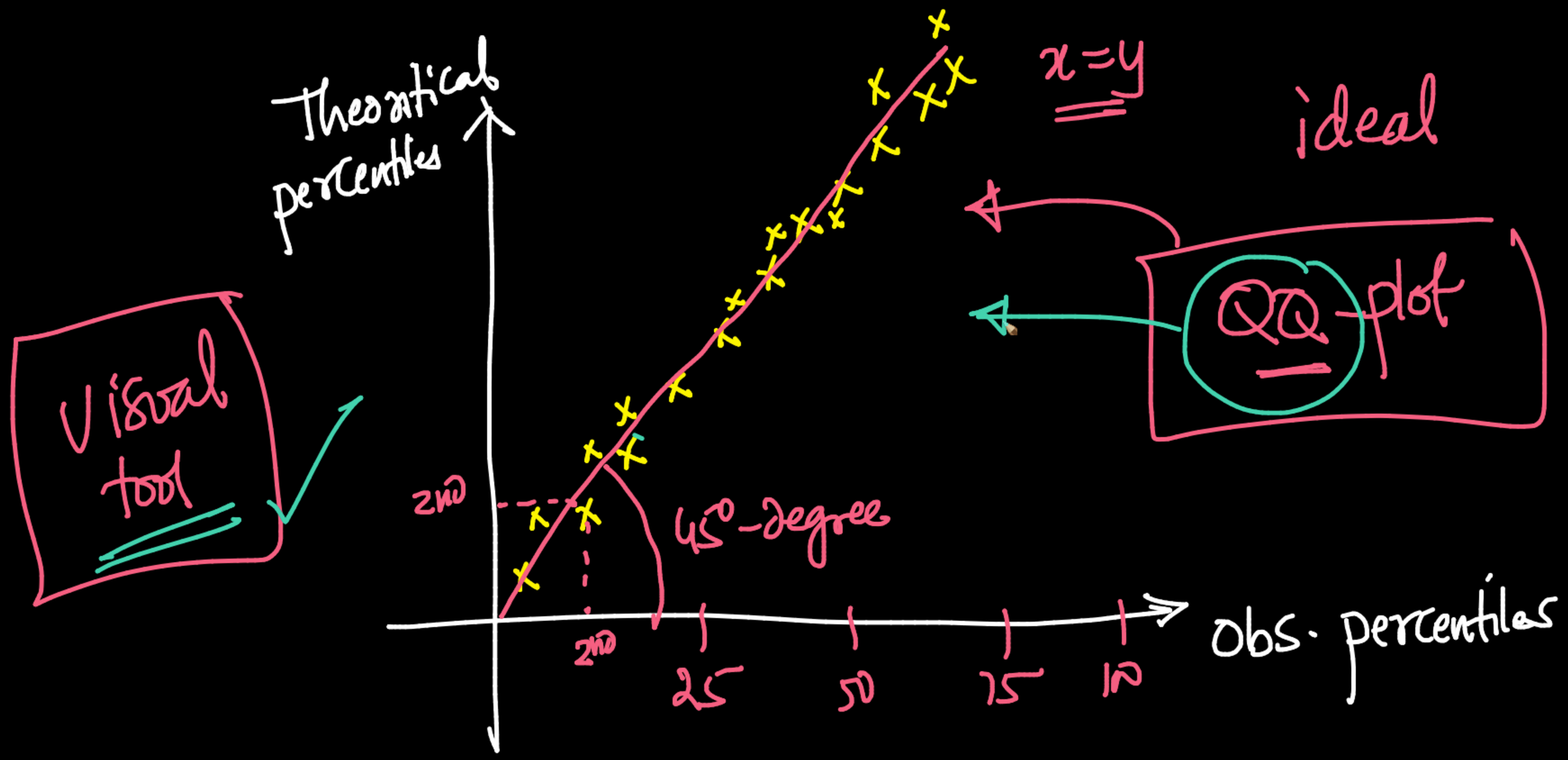
$N(\mu=160, \sigma=10)$

plot



62 / 63

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:30



63 / 64

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:30

$x_1, \ldots x_{100} \longrightarrow \text{Normal}(\mu, \sigma)$

theoretical percentiles

obs-percentiles

Dis-adv

Can't Quantify (as a prob) how confident you are

hyp-testing

Close to

x=y line

45°-degree

Very confident

64 / 64

10-15 min →

5 lines

Code → next class

QQplot

sample

$x_1, x_2 \ldots x_{50}$

90 min →

Conf. Intervals → Bootstrapping

→ CLT

pop-mean

↓

[,]

pop-median

✓ Hyp-testing & Exp-design ← one of the hardest concepts

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:41

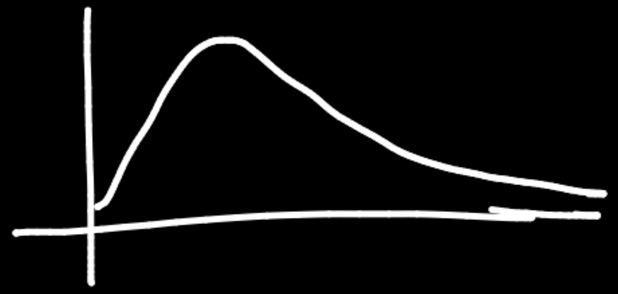
65 / 68

$x_1, x_2, \ldots x_{100}$

Conditions X

① PDF → shape (wiki)

CDF



② Guess: log-normal

$\log(x_i)$'s → Normal(μ, σ)

$\mu \to \bar{x}$, $\sigma \to s$

QQ-plot

empirical $\bar{x}$, s

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:36

→ Guess! . . .

Another → $x_1, x_2, \ldots x_n$

Transformation

box-cox Transform

$\boxed{f}(x_i)$

[NOT-ALWAYS-POSSIBLE]

x_i's ~ Normal . . .

67 / 67

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:41

68 / 68