iii A MATHEMATICALLY MORE CONVENIENT WAY TO WRITE THE MODEL

$$P_X(x) = \mu^{x}(1-\mu)^{1-x}, \quad x \in \{0,1\}, \quad \mu \in (0,1)$$

BERNOULLI PROBABILITY MASS FUNCTION

$$p_x(0) = \mu^0 (1-\mu)^{1-0} = 1-\mu$$

$$p_x(1) = \mu^1 (1-\mu)^0 = \mu$$

FROM MODELS TO DATA &

FROM PROBABILITIES TO FREQUENCIES

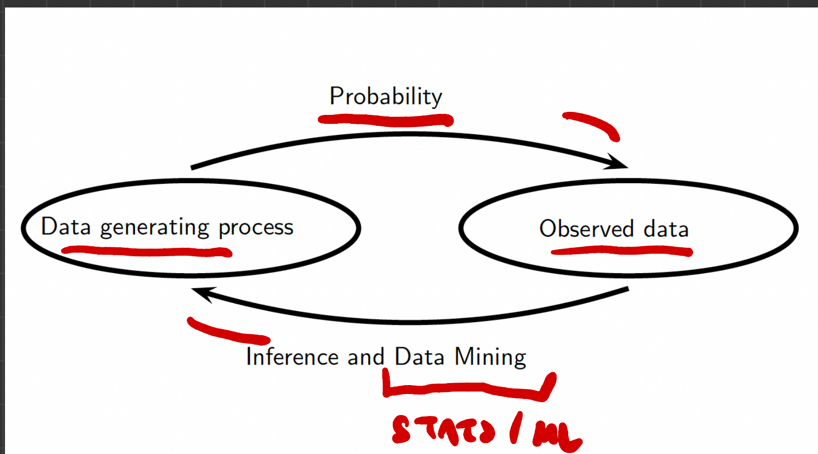
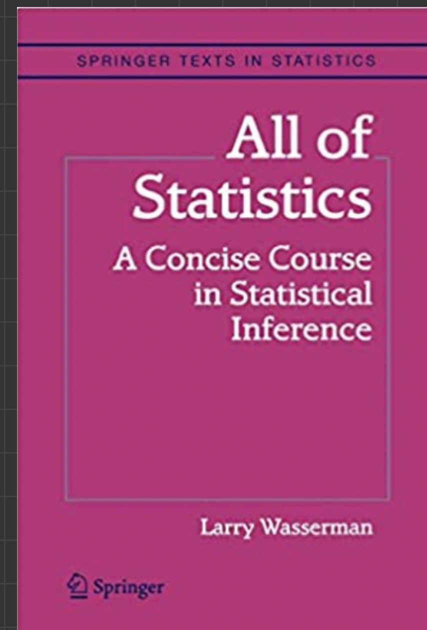


FIGURE 1. Probability and inference.



- LET'S ASSUME THAT OBSERVED DATA ARE GENERATED FROM BERNOULLI MODEL

$$X_i \sim \mathrm{Bern}(\mu) \qquad i = 1, \dots, n$$

10:51 Fri 20 Feb

New tab | Bemacs_course - Posit

posit.cloud

The ability to publish applications and documents to Posit Cloud is now removed for all users. Please visit **Posit Connect Cloud** to support your deployment needs.

posit Cloud

Your Workspace / Bemacs_course

Om P

Spaces
- Your Workspace
- New Space

Learn
- Guide
- What's New
- Recipes
- Cheatsheets

Help
- Current System Status
- Posit Community

Info
- Plans & Pricing
- Terms and Conditions

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | R 4.1.3

Source

Console | Terminal | Background Jobs

R 4.1.3 · /cloud/project/

```
Session restored from your saved work on 2026-Feb-1
3 19:22:21 UTC (6 days ago)
> x=rbinom(10,1,0.8)
> x
 [1] 0 1 1 1 0 1 1 1 1 1
> ?binom
No documentation for ‘binom’ in specified packages
and libraries:
you could try ‘??binom’
> ??binom
> x=rbinom(10,1,0.8)
> x
 [1] 0 1 0 0 1 1 1 1 1 1
> y=rbinom(10,1,0.2)
> y
 [1] 0 0 1 0 0 0 0 0 0 0
>
```

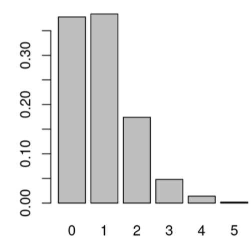
Environment | Help

R: Search Results | Find in Topic

base::Special	Special Functions of Mathematics
MASS::anova.negbin	Likelihood Ratio Tests for Negative Binomial GLMs
MASS::dose.p	Predict Doses for Binomial Assay model
MASS::glm.convert	Change a Negative Binomial fit to a GLM fit
MASS::glm.nb	Fit a Negative Binomial Generalized Linear Model

Plots | Viewer | Presentation

Zoom | Export



PROBABILITIES & FREQUENCIES

$X_1, \ldots, X_n$ BINARY RV $X_i \sim \mathrm{Bern}(\mu)$

$Y_n = \sum_{i=1}^{n} X_i$ FREQ. OF 1's IN THE SAMPLE

$\frac{1}{n} Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ PERC. " "

10:54 Fri 20 Feb ••• 77%

New tab | Bemacs_course - Posit

posit.cloud

The ability to publish applications and documents to Posit Cloud is now removed for all users. Please visit **Posit Connect Cloud** to support your deployment needs.

posit Cloud

Your Workspace / Bemacs_course

Om P

Spaces
- Your Workspace
- New Space

Learn
- Guide
- What's New
- Recipes
- Cheatsheets

Help
- Current System Status
- Posit Community

Info
- Plans & Pricing
- Terms and Conditions

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | R 4.1.3

Source

Console | Terminal | Background Jobs

R 4.1.3 · /cloud/project/

```
 [1] 0 0 1 0 0 0 0 0 0 0
> table(x)
x
0 1
3 7
> x=rbinom(100,1,0.8)
> table(x)
x
 0  1
14 86
> x
  [1] 1 1 1 0 1 1 1 1 1 0 1 1 1 1 0 1 1 1 1 1 1
 [22] 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1
 [43] 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1
 [64] 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 0 1 1 1 1 1
 [85] 1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 0
> x=rbinom(1000,1,0.8)
> table(
x
  0   1
209 791
>
```

chrome://newtab | chrome://newtab and 1 other | posit.cloud and 1 other | New Window

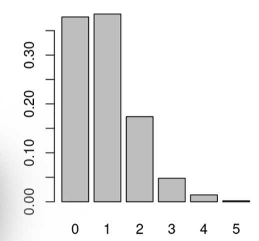
Environment | Help

R: Search Results | Find in Topic

base::Special	Special Functions of Mathematics
MASS::anova.negbin	Likelihood Ratio Tests for Negative Binomial GLMs
MASS::dose.p	Predict Doses for Binomial Assay model
MASS::glm.convert	Change a Negative Binomial fit to a GLM fit
MASS::glm.nb	Fit a Negative Binomial Generalized Linear Model

Plots | Viewer | Presentation

Zoom | Export



OBSERVATION FROM SIMULATIONS

As n gets larger

$$\frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{concentrates around } \mu$$

↳ there "concentration"

MAXIMUM LIKELIHOOD OR MINIMUM ENTROPY ESTIMATION

THIS GENERATED THE FOLLOWING IDEA:

GIVEN BINARY DATA: $x_1, \ldots, x_n$

THEN I COULD USE THE

SAMPLE PERCENTAGE OF 1'S TO

LEARN A PROBABILITY MODEL FROM

THESE DATA

LEARN THE BERNOULLI PARAMETER

MODELLING COUNT DATA

→ RECALL : COUNT DATA RELATE TO QUESTION
" HOW MANY " (ESPECIALLY IF RARE)

→ GOALS IN FOOTBALL GAME
→ INSURANCE CLAIMS
→ WRONGFUL CONVICTIONS PER YEAR
19'TH CENTURY FRANCE

↳ POISSON

THE POISSON MODEL / DISTRIBUTION

WITH PROBABILITY MASS FUNCTION

$$P_X(x) = \frac{e^{-\mu} \mu^{x}}{x!} \qquad x \in \{0, 1, 2, \ldots\}$$

$$\mu > 0$$

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

$e = 2.71\ldots$ EULER CONSTANT

$x! = x(x-1)\cdots 1$ $x \in \{1, 2, \ldots\}$

FACTORIAL

$3! = 3 \cdot 2 \cdot 1 = 6$

BY CONVENTION $0! = 1$

i. $p_X(x) \geq 0$

$$\left.\begin{array}{ll} e^{-\mu} \geq 0 & \\ \mu^{x} \geq 0 & \mu > 0 \\ x! \geq 0 & x \in \{0, 1, \dots\} \end{array}\right\} \frac{e^{-\mu}\mu^{x}}{x!} \geq 0$$

FAR LESS OBVIOUS IS $\dfrac{e^{-\mu}\mu^{x}}{x!} \leq 1$

ii $P_X(0) + P_X(1) + \cdots = \sum_{x=0}^{\infty} P_X(x) = 1$

$$\boxed{\begin{aligned} \sum_{i=0}^{\infty} a_i &= \lim_{n\to\infty} \sum_{i=0}^{n} a_i \\ \sum_{i=0}^{n} a_i &= a_0 + a_1 + \cdots + a_n \end{aligned}}$$

REMINDER

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

$$= \sum_{x=0}^{\infty} \frac{1}{x!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \dots$$

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

POWER SERIES EXPANSION

SINCE $$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$ THEN

$$1 = e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^x}{x!} = \sum_{x=0}^{\infty} P_X(x)$$

$$\Rightarrow \quad P_X(x) \leq 1$$

ge
s
events
article
Vikipedia
us
ute
edit
nity portal
changes
file
ks here
changes
pages
ent link

Poisson distribution

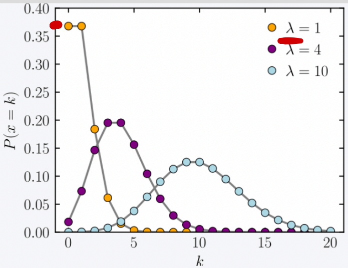
From Wikipedia, the free encyclopedia

In probability theory and statistics, the **Poisson distribution** (/ˈpwɑːsɒn/; French pronunciation: [pwasɔ̃]), named after French mathematician Siméon Denis Poisson, 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] 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

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.

$k \leftarrow x$

$P(x = k)$

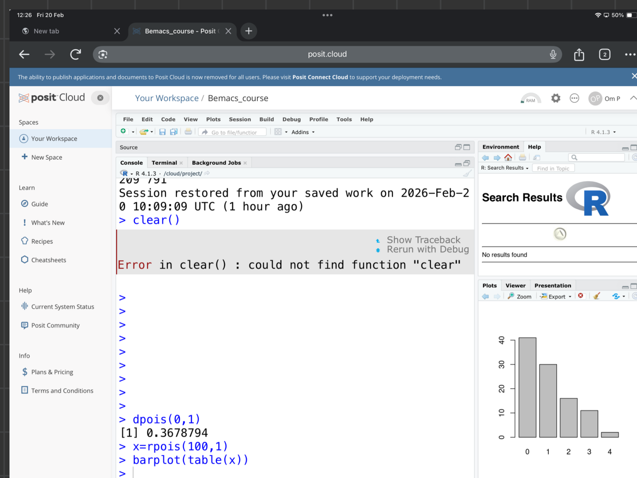
$P_x(x)$

$\lambda \leftarrow \mu$

Numerical experiments

- Generate $X_i \sim \text{Poisson}(1)$

$\rightarrow p_X(0) \simeq 0.37$



1st OBSERVATION

PERCENTAGE OF DATA IN SAMPLE GENERATED FROM THE MODEL EQUAL TO x

CONCENTRATED AROUND $p_X(x)$

↳ DERIVE FROM RESULT FOR BERNOULLI

ANOTHER OBSERVATION

- $\frac{1}{n}\sum_{i=1}^{n} X_i$ CONCENTRATES AROUND μ

  ↳ IDEA FOR LEARNING A POISSON MODEL FROM OBSERVED DATA

  ↳ μ BY $\frac{1}{n}\sum_{i=1}^{n} x_i$