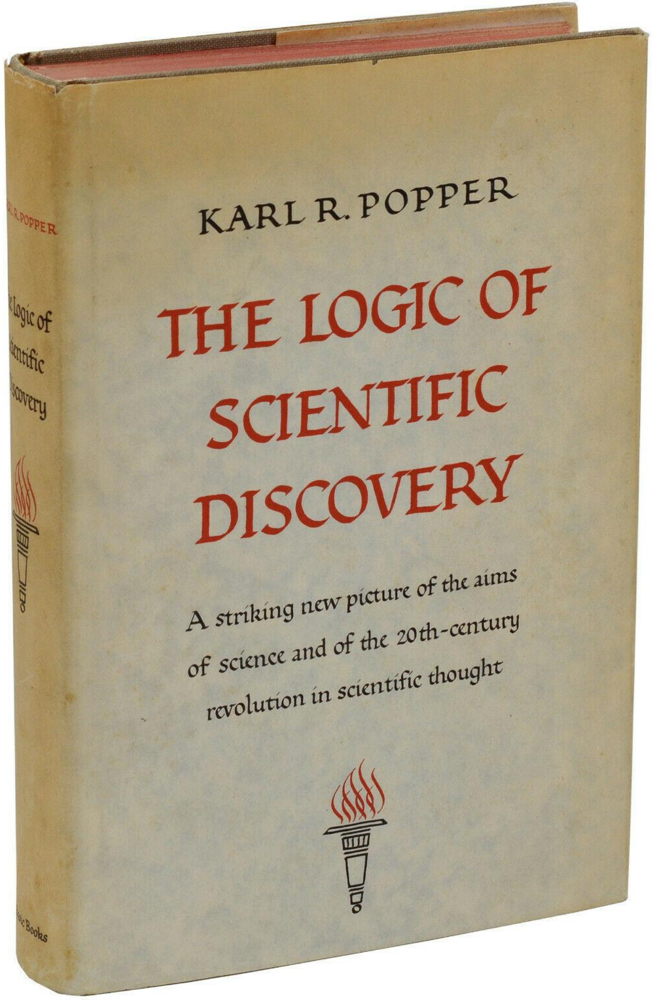
My proposal is based upon an *asymmetry* between **verifiability** and **falsifiability**; an asymmetry which results from the logical form of universal statements. For these are never derivable from singular statements, but can be contradicted by singular statements.
—Karl Popper, *The Logic of Scientific Discovery*

What is science?

a scientific theory must be *falsifiable*



Reproducibility

Reproducible research means:

the ability of a researcher to duplicate the results of a prior study using the same materials as were used by the original investigator. That is, a second researcher might use the same raw data to build the same analysis files and implement the same statistical analysis in an attempt to yield the same results

https://acmedsci.ac.uk/viewFile/56314e40aac61.pdf

Reproducible research in practice:

all numbers in a data analysis can be recalculated exactly (down to stochastic variables!) using the **code** and **raw data** provided by the analyst.

- provide raw data and code to reduce it to all stages needed to get outputs
- provide code to reproduce all figures
- provide code to reproduce all number outcomes

Probability

Frequentist interpretation

fraction of times something happens

probability of it happening

Bayesian interpretation

represents a level of certainty relating to a potential outcome or idea:

if I believe the coin is unfair (tricked) then even if I get a head and a tail I will still believe I am more likely to get heads than tails

Frequentist interpretation

fraction of times something happens

$$\Updownarrow$$

probability of it happening

P(E) = frequency of E

P(coin = head) = 6/11 = 0.55

Frequentist interpretation

fraction of times something happens

$$\Updownarrow$$

probability of it happening

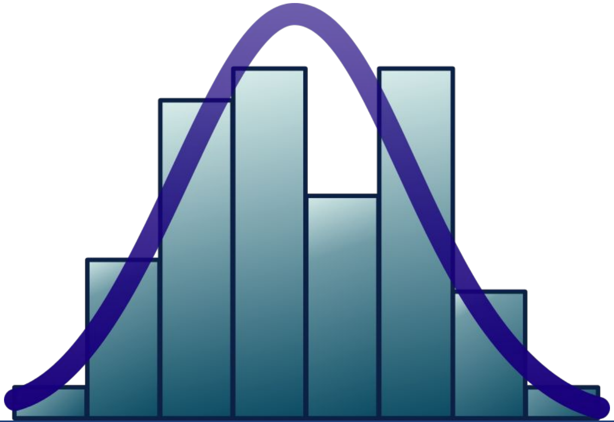
$$P(E) = \text{frequency of } E$$

$$P(\text{coin} = \text{head}) = 6/11 = 0.55$$

$$P(\text{coin} = \text{head}) = 49/100 = 0.49$$

Summary of probabilities

Event	Probability
A	$P(A) \in [0,1]$
not A	$P(A^\complement) = 1 - P(A)$
A or B	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $P(A \cup B) = P(A) + P(B) \qquad \text{if A and B are mutually exclusive}$
A and B	$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$ $P(A \cap B) = P(A)P(B) \qquad \text{if A and B are independent}$
A given B	$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \mid A)P(A)}{P(B)}$



Statistics

Statistics

takes us from observing a limited number of samples to infer on the population

Taxonomy

Distribution: a formula (a model)

Population: all of the elements of a "family"

Sample: a finite subset of the population that you observe

Descriptive statistics - describing the central tendency

mean: (sum of all the terms)/(number of terms)

median: 50% of the distribution is to the left,

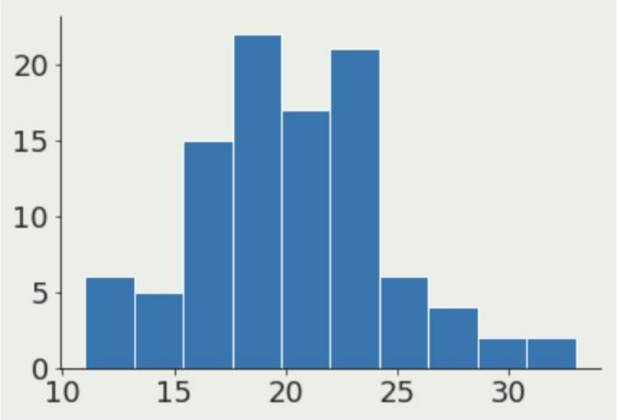
50% to the right

mode: most popular value in the distribution

```
dist = sp.stats.poisson.rvs(size=100, mu=20)
pl.hist(dist)
print(dist.mean())
print(np.median(dist))
print(sp.stats.mode(dist))
```

executed in 125ms, finished 15:01:20 2019-09-09

```
20.06
20.0
ModeResult(mode=array([18]), count=array([12]))
```



Taxonomy

Central tendency: mean, median, mode

Spread: variance, interquartile range

Descriptive statistics - measuring the spread

Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right]$$

Standard deviation

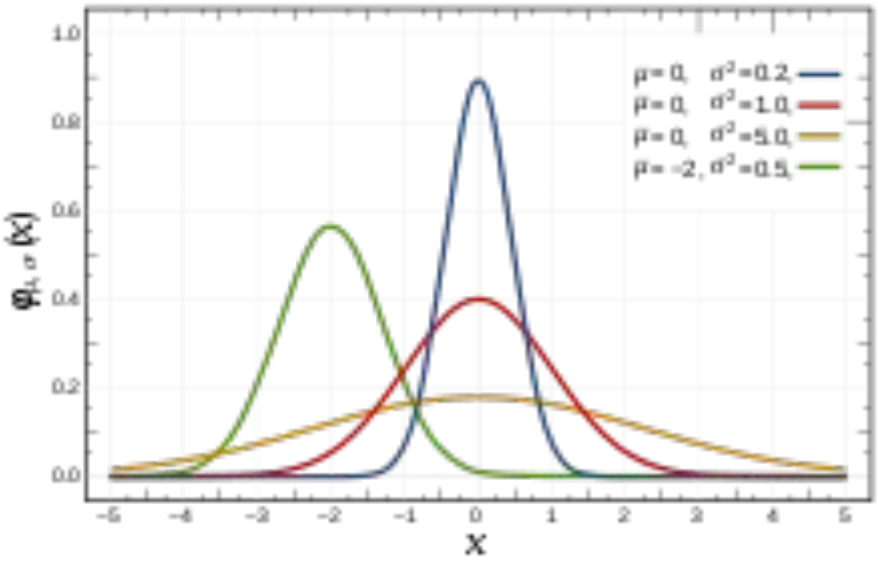
$$\sigma(X) = \sqrt{\mathrm{E}\left[(X-\mathrm{E}[X])^2\right]}$$

Gaussian distribution:

1σ contains 68% of the distribution

2σ contains 95% of the distribution

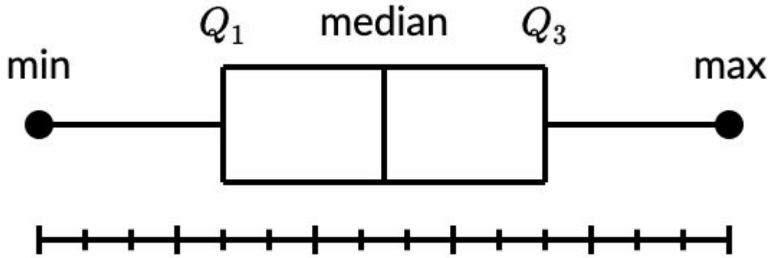
3σ contains 97.3% of the distribution



Descriptive statistics - measuring the spread

Interquartile range

Where are the limits within which X% of the distribution is contained



Probability distributions

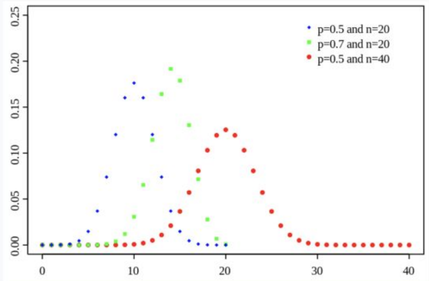
Binomial

Coin toss

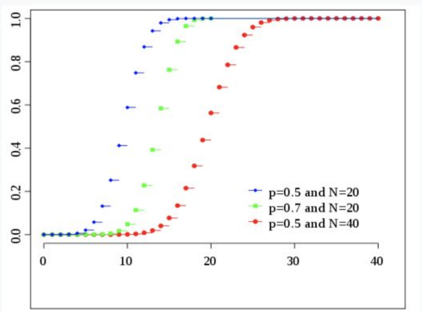
Fair coin = p=0.5 n=1

Binomial distribution

Probability mass function



Cumulative distribution function



Notation	$B(n,p)$
Parameters	$n \in \{0,1,2,\ldots\}$ – number of trials $p \in [0,1]$ – success probability for each trial
Support	$k \in \{0,1,\ldots,n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Notation	$B(n,p)$
Parameters	$n \in \{0,1,2,\ldots\}$ – number of trials $p \in [0,1]$ – success probability for each trial
Support	$k \in \{0,1,\ldots,n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Variance	$np(1-p)$
Skewness	$\frac{1-2p}{\sqrt{np(1-p)}}$
Ex. kurtosis	$\frac{1-6p(1-p)}{np(1-p)}$
Entropy	$\frac{1}{2}\log_2(2\pi enp(1-p)) + O\left(\frac{1}{n}\right)$ in shannons. For nats, use the natural log in the log.
MGF	$(1-p+pe^t)^n$
CF	$(1-p+pe^{it})^n$
PGF	$G(z) = [(1-p)+pz]^n$
Fisher information	$g_n(p) = \frac{n}{p(1-p)}$ (for fixed n)

Probability distributions

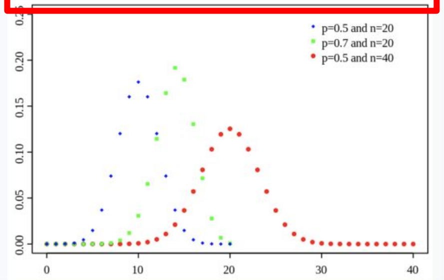
Binomial

Coin toss

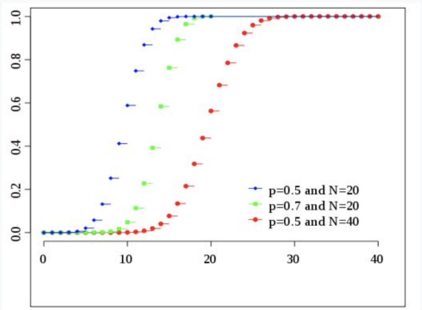
Fair coin = p=0.5 n=1

Binomial distribution

Probability mass function



Cumulative distribution function



Notation	$B(n, p)$
Parameters	$n \in \{0, 1, 2, \ldots\}$ – number of trials $p \in [0, 1]$ – success probability for each trial
Support	$k \in \{0, 1, \ldots, n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Notation	$B(n, p)$
Parameters	$n \in \{0, 1, 2, \ldots\}$ – number of trials $p \in [0, 1]$ – success probability for each trial
Support	$k \in \{0, 1, \ldots, n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$
Variance	$np(1-p)$
Skewness	$\frac{1-2p}{\sqrt{np(1-p)}}$
Ex. kurtosis	$\frac{1-6p(1-p)}{np(1-p)}$
Entropy	$\frac{1}{2}\log_2(2\pi enp(1-p)) + O\left(\frac{1}{n}\right)$ in shannons. For nats, use the natural log in the log.
MGF	$(1-p+pe^t)^n$
CF	$(1-p+pe^{it})^n$
PGF	$G(z) = [(1-p)+pz]^n$
Fisher information	$g_n(p) = \frac{n}{p(1-p)}$ (for fixed n)

Probability distributions

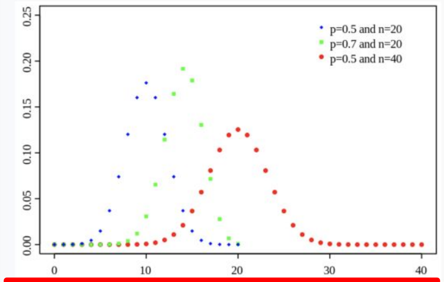
Binomial

Coin toss

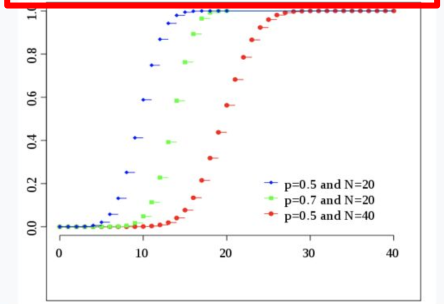
Fair coin = p=0.5 n=1

Binomial distribution

Probability mass function



Cumulative distribution function



Notation	$B(n,p)$
Parameters	$n \in \{0,1,2,\ldots\}$ – number of trials $p \in [0,1]$ – success probability for each trial
Support	$k \in \{0,1,\ldots,n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Notation	$B(n,p)$
Parameters	$n \in \{0,1,2,\ldots\}$ – number of trials $p \in [0,1]$ – success probability for each trial
Support	$k \in \{0,1,\ldots,n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$
Variance	$np(1-p)$
Skewness	$\frac{1-2p}{\sqrt{np(1-p)}}$
Ex. kurtosis	$\frac{1-6p(1-p)}{np(1-p)}$
Entropy	$\frac{1}{2}\log_2(2\pi e n p(1-p)) + O\left(\frac{1}{n}\right)$ in shannons. For nats, use the natural log in the log.
MGF	$(1-p+pe^t)^n$
CF	$(1-p+pe^{it})^n$
PGF	$G(z) = [(1-p)+pz]^n$
Fisher information	$g_n(p) = \frac{n}{p(1-p)}$ (for fixed n)

Probability distributions

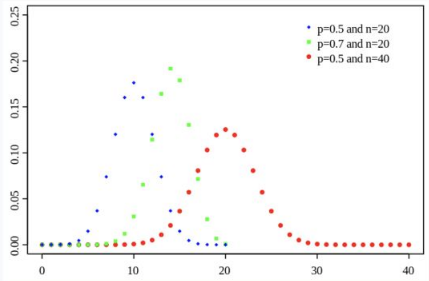
Binomial

Coin toss

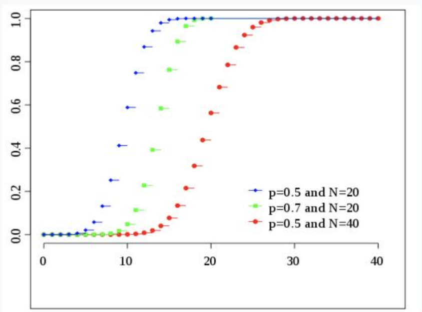
Fair coin = p=0.5 n=1

Binomial distribution

Probability mass function



Cumulative distribution function



Notation	$B(n,p)$
Parameters	$n \in \{0,1,2,\ldots\}$ – number of trials $p \in [0,1]$ – success probability for each trial
Support	$k \in \{0,1,\ldots,n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Notation	$B(n,p)$
Parameters	$n \in \{0,1,2,\ldots\}$ – number of trials $p \in [0,1]$ – success probability for each trial
Support	$k \in \{0,1,\ldots,n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$
Variance	$np(1-p)$
Skewness	$\frac{1-2p}{\sqrt{np(1-p)}}$
Ex. kurtosis	$\frac{1-6p(1-p)}{np(1-p)}$
Entropy	$\frac{1}{2}\log_2(2\pi enp(1-p)) + O\left(\frac{1}{n}\right)$ in shannons. For nats, use the natural log in the log.
MGF	$(1-p+pe^t)^n$
CF	$(1-p+pe^{it})^n$
PGF	$G(z) = [(1-p)+pz]^n$
Fisher information	$g_n(p) = \frac{n}{p(1-p)}$ (for fixed n)

Probability distributions

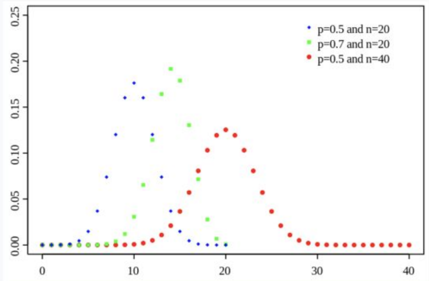
Binomial

Coin toss

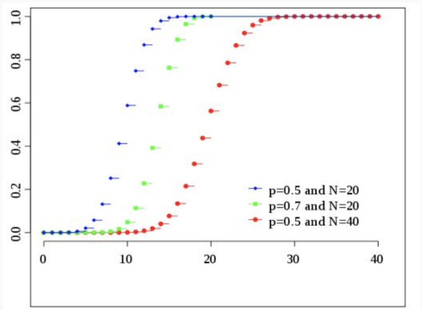
Fair coin = p=0.5 n=1

Binomial distribution

Probability mass function



Cumulative distribution function



Notation	$B(n, p)$
Parameters	$n \in \{0, 1, 2, \ldots\}$ – number of trials $p \in [0, 1]$ – success probability for each trial
Support	$k \in \{0, 1, \ldots, n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Notation	$B(n, p)$
Parameters	$n \in \{0, 1, 2, \ldots\}$ – number of trials $p \in [0, 1]$ – success probability for each trial
Support	$k \in \{0, 1, \ldots, n\}$ – number of successes
pmf	$\binom{n}{k} p^k (1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$
Mean	np
Median	$\lfloor np \rfloor$ or $\lceil np \rceil$
Mode	$\lfloor (n+1)p \rfloor$ or $\lceil (n+1)p \rceil - 1$

Variance	$np(1-p)$
Skewness	$\dfrac{1-2p}{\sqrt{np(1-p)}}$
Ex. kurtosis	$\dfrac{1-6p(1-p)}{np(1-p)}$
Entropy	$\frac{1}{2}\log_2(2\pi enp(1-p)) + O\left(\frac{1}{n}\right)$ in shannons. For nats, use the natural log in the log.
MGF	$(1-p+pe^t)^n$
CF	$(1-p+pe^{it})^n$
PGF	$G(z) = [(1-p)+pz]^n$
Fisher information	$g_n(p) = \dfrac{n}{p(1-p)}$ (for fixed n)

ETH zürich

Probability distributions

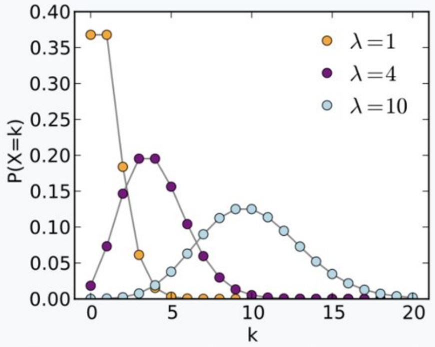
Poisson

Shut noise/count noise

The innate noise in natural steady state processes (star flux, rain drops, radioactive decay...)

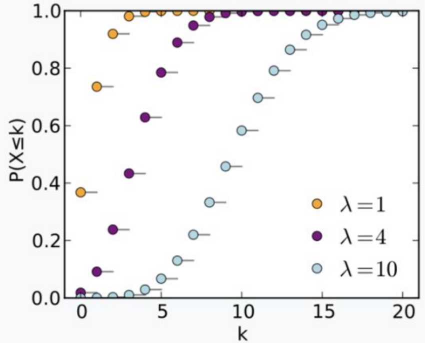
Poisson

Probability mass function



The horizontal axis is the index *k*, the number of occurrences. *λ* is the expected number of occurrences, which need not be an integer. 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	$\mathrm{Pois}(\lambda)$
Parameters	$\lambda > 0$, (real) — rate
Support	$k \in \{0, 1, 2, \ldots\}$
pmf	$\frac{\lambda^k e^{-\lambda}}{k!}$
CDF	$\frac{\Gamma(\lfloor k+1 \rfloor, \lambda)}{\lfloor k \rfloor !}$, or $e^{-\lambda} \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}$, 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
Skewness	$\lambda^{-1/2}$
Ex. kurtosis	λ^{-1}
Entropy	$\lambda[1 - \log(\lambda)] + e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k \log(k!)}{k!}$ (for large λ) $\frac{1}{2}\log(2\pi e \lambda) - \frac{1}{12\lambda} - \frac{1}{24\lambda^2} - \frac{19}{360\lambda^3} + O\left(\frac{1}{\lambda^4}\right)$
MGF	$\exp(\lambda(e^t - 1))$
CF	$\exp(\lambda(e^{it} - 1))$
PGF	$\exp(\lambda(z - 1))$
Fisher information	$\frac{1}{\lambda}$

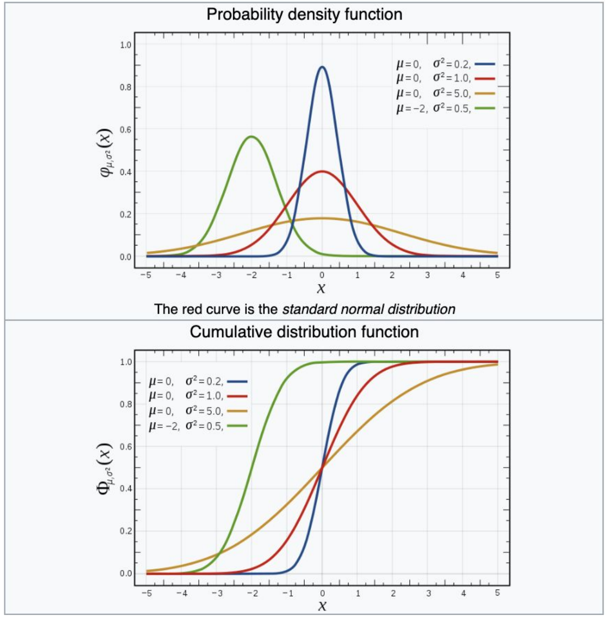
Probability distributions

Gaussian

most common noise:

well behaved mathematically, symmetric, when we can we will assume our uncertainties are Gaussian distributed

Normal Distribution



Notation	$\mathcal{N}(\mu, \sigma^2)$
Parameters	$\mu \in \mathbb{R}$ = mean (location) $\sigma^2 > 0$ = variance (squared scale)
Support	$x \in \mathbb{R}$
PDF	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
CDF	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$
Quantile	$\mu + \sigma\sqrt{2}\operatorname{erf}^{-1}(2F-1)$
Mean	μ
Median	μ
Mode	μ
Variance	σ^2
Skewness	0
Ex. kurtosis	0
Entropy	$\frac{1}{2}\log(2\pi e\sigma^2)$
MGF	$\exp(\mu t + \sigma^2 t^2/2)$
CF	$\exp(i\mu t - \sigma^2 t^2/2)$
Fisher information	$\mathcal{I}(\mu,\sigma) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 2/\sigma^2 \end{pmatrix}$ $\mathcal{I}(\mu,\sigma^2) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(2\sigma^4) \end{pmatrix}$
Kullback-Leibler divergence	$D_{\text{KL}}(\mathcal{N}_0 \\| \mathcal{N}_1) = \frac{1}{2}\{(\sigma_0/\sigma_1)^2 + \frac{(\mu_1-\mu_0)^2}{\sigma_1^2} - 1 + 2\ln\frac{\sigma_1}{\sigma_0}\}$

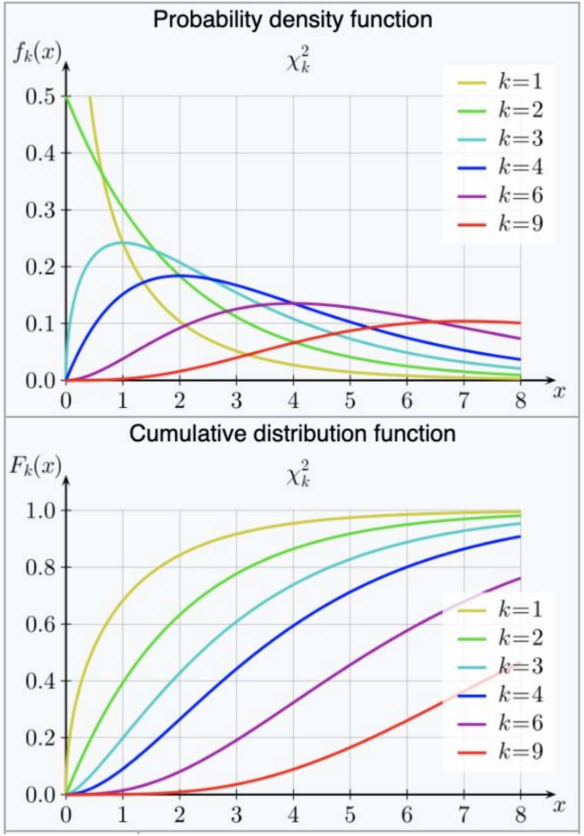
Probability distributions

Chi-square (X2)

turns out its extremely common

many pivotal quantities follow this distribution and thus many tests are based on this

chi-square



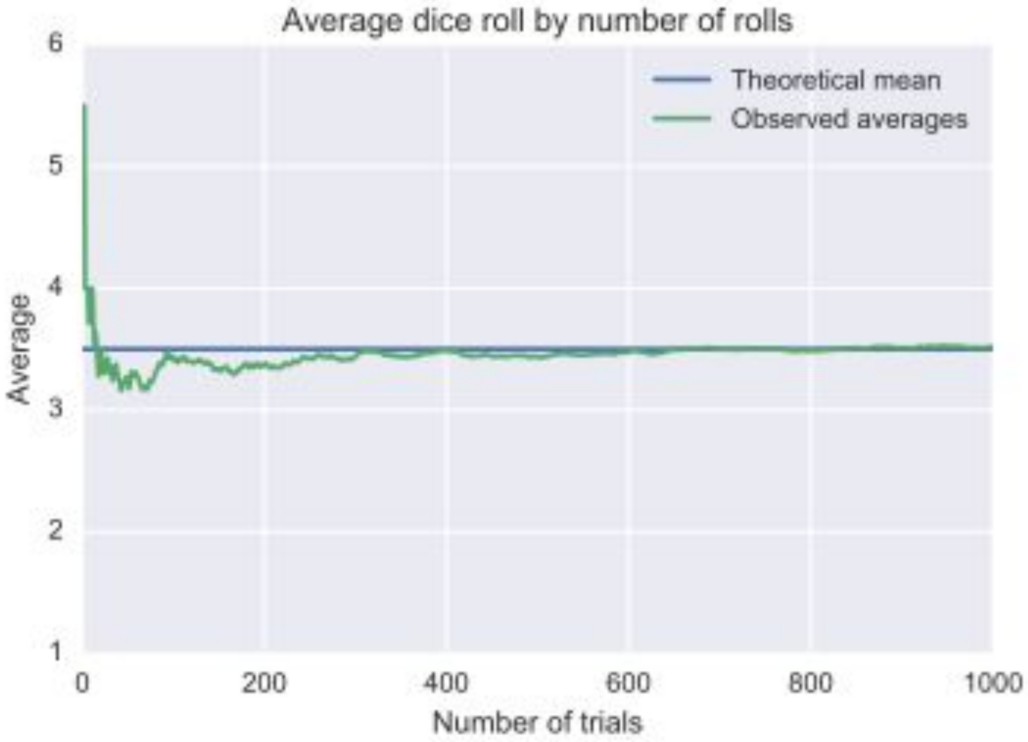
Notation	$\chi^2(k)$ or χ^2_k
Parameters	$k \in \mathbb{N}_{>0}$ (known as "degrees of freedom")
Support	$x \in (0, +\infty)$ if $k = 1$, otherwise $x \in [0, +\infty)$
PDF	$\frac{1}{2^{k/2}\Gamma(k/2)}\, x^{k/2-1} e^{-x/2}$
CDF	$\frac{1}{\Gamma(k/2)}\, \gamma\left(\frac{k}{2}, \frac{x}{2}\right)$
Mean	k
Median	$\approx k\left(1 - \frac{2}{9k}\right)^3$
Mode	$\max(k-2, 0)$
Variance	$2k$
Skewness	$\sqrt{8/k}$
Ex. kurtosis	$\frac{12}{k}$
Entropy	$\frac{k}{2} + \ln(2\Gamma(\frac{k}{2})) + (1 - \frac{k}{2})\psi(\frac{k}{2})$
MGF	$(1-2t)^{-k/2}$ for $t < \frac{1}{2}$
CF	$(1-2it)^{-k/2}$ [1]
PGF	$(1 - 2\ln t)^{-k/2}$ for $0 < t < \sqrt{e}$

Simulation of the different distributions in Colab

Law of large numbers

According to the law, the average of the results obtained from a large number of trials should be close to the expected value and will tend to become closer to the expected value as more trials are performed.

In the limit of N -> infinity the mean of a sample of size N approaches the mean of the population μ

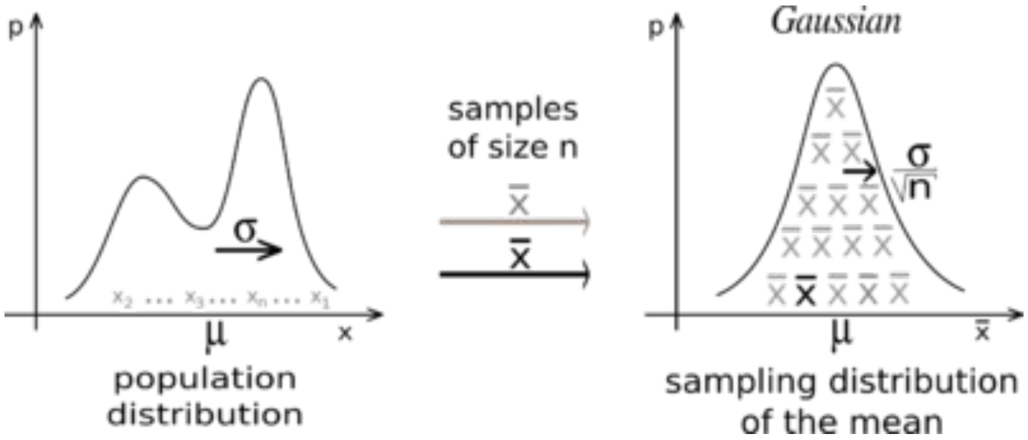


Central limit theorem

In the limit of N -> infinity

the sample mean x approaches a Normal (Gaussian) distribution with mean μ and standard deviation σ

regardless of the distribution of X



Central limit theorem in Colab

Inference

"Body weight was higher in mice fed the high-fat diet already after the first week, due to higher dietary intake in combination with lower metabolic efficiency."

"Already during the first week after introduction of high-fat diet, body weight increased significantly more in the high-fat diet-fed mice (+ 1.6 ± 0.1 g) than in the normal diet-fed mice (+ 0.2 ± 0.1 g; P < 0.001)."

What is + 1.6 ± 0.1g and + 0.2 ± 0.1 g and what is P < 0.001?
And why are these numbers important?

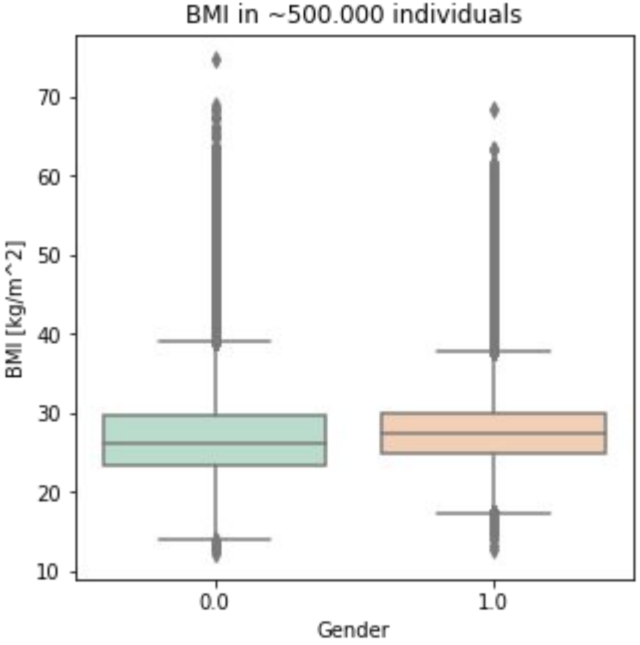
Data has variability

We need p-values and confidence intervals because the data has variability.

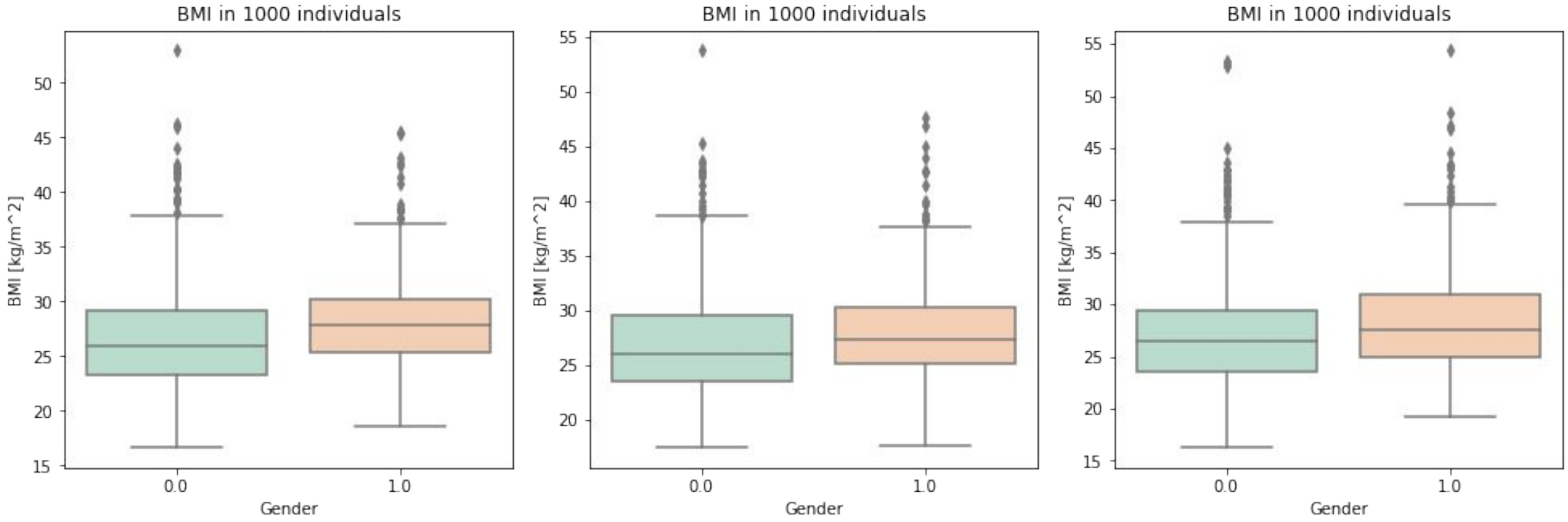
We do not have access to the ***population***, we are just sampling it and inferring the characteristics of the population based on the values of its sample.

Every time we resample the population, we get different descriptive statistics (mean, standard deviation, etc).

In particular, the means is a random variable.



Random variables when sampling a population




```
mean males:  28.065288965238878
mean females:  26.76744359319031
```

```
mean males:  27.9795995248824
mean females:  26.92432748169731
```

```
mean males:  28.379904688459526
mean females:  27.187517814945377
```

The Null Hypothesis

How do we know that males and females have really different BMI, and that is not due to a random effect of the sampling?

Null Hypothesis Rejection Testing

Null Hypothesis Rejection Testing

model → prediction

data → does not falsify → model still holds

data → falsifies → model rejected

Null Hypothesis Rejection Testing

model → prediction

"Under the hypothesis" =
If the model is true

This happens

data → does not falsify → model still holds

data → falsifies → model rejected

ETH zürich

Null Hypothesis Rejection Testing

model → prediction

"Under the **NULL** hypothesis" =
If the NH model is true

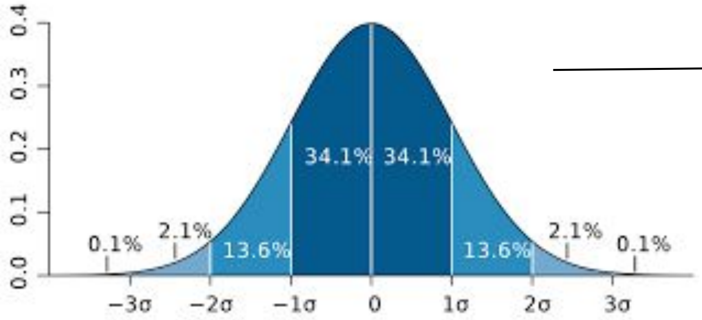
This has **low** probability of happening

data → does not falsify → model still holds

data → falsifies → model rejected

low probability event happened

Null Hypothesis Rejection Testing



2σ

rejected at 95%

0.05 p-value

5% confidence

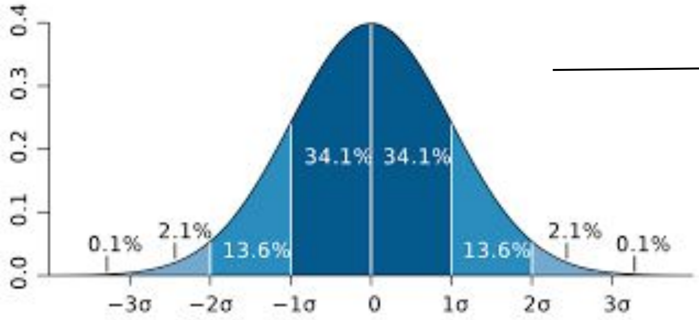
data

does not falsify → model still holds

falsifies → model rejected

low probability event happened

Null Hypothesis Rejection Testing



3σ

rejected at 99.7%

0.003 p-value

0.3% confidence

data

does not falsify

model

still holds

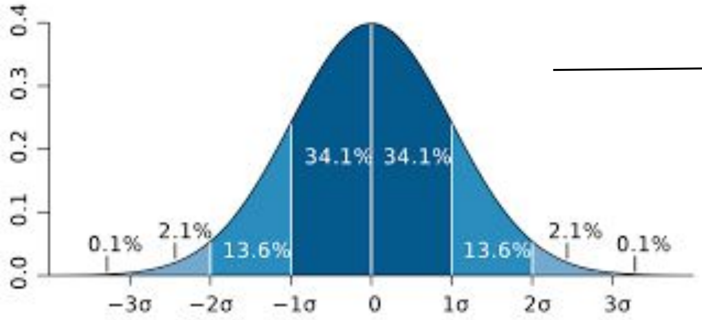
falsifies

model

rejected

low probability event happened

Null Hypothesis Rejection Testing



5σ

rejected at 99.99..%

3-e7 p-value

3-e5% confidence

data

does not falsify

model

still holds

falsifies

low probability event happened

model

rejected

Null Hypothesis Rejection Testing

formulate the Null as the comprehensive opposite of your theory

model → prediction

“Under the **NULL** hypothesis” = If the proposed model is **FALSE**

This has **low** probability of happening

data → does not falsify alternative → model rejected

data → Falsifies alternative → model still holds

low probability event happened

Example

Earth is flat is Null

Earth is ~~round~~ not flat is Alternative

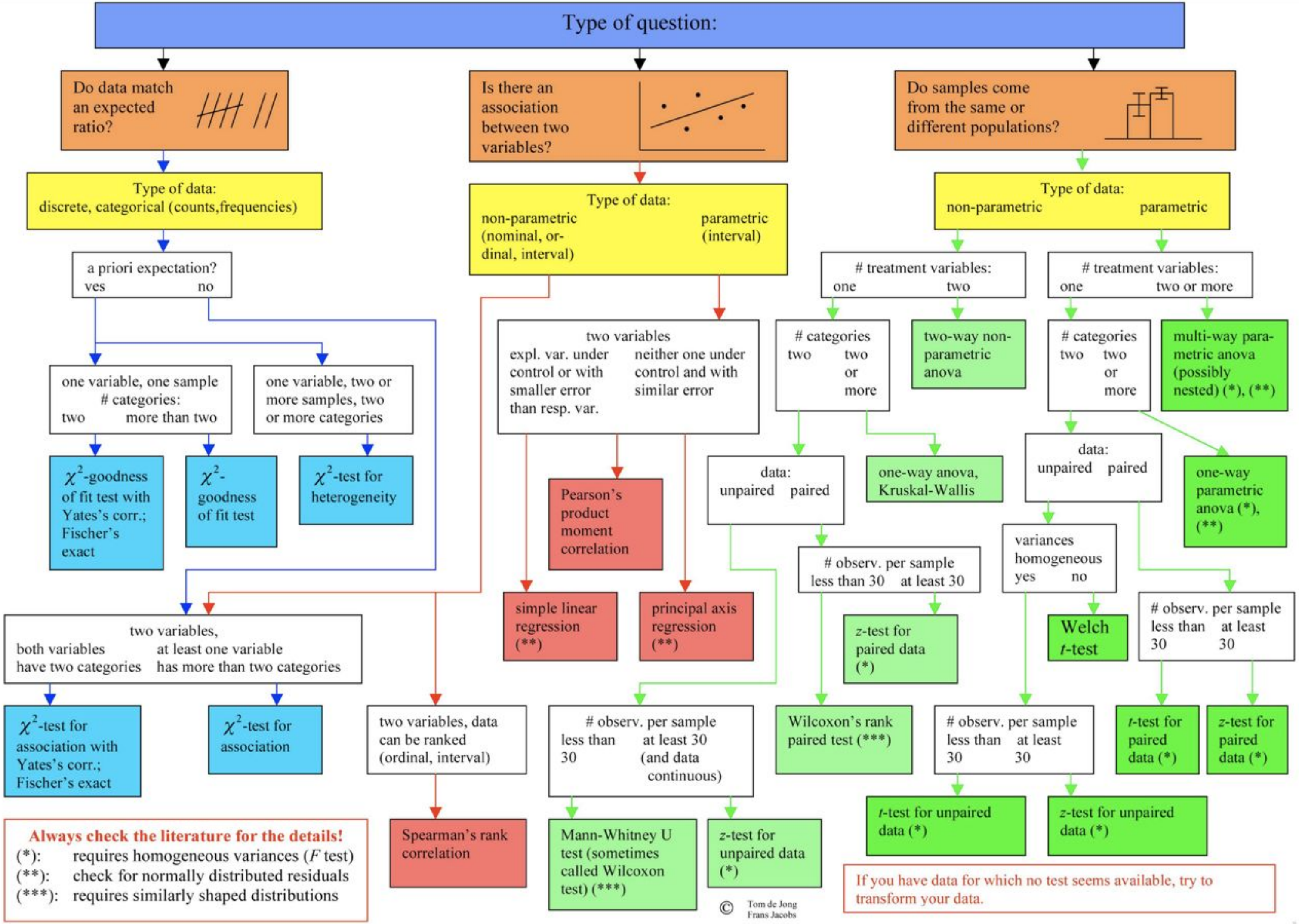
We reject the Null hypothesis that the Earth is flat (p=0.05)

Null Hypothesis Rejection Testing

1. Formulate your prediction (Null Hypothesis)
2. Identify all alternative outcomes (Alternative hypothesis):
   a. If all alternatives to our model are ruled out, then our model must hold
3. Set confidence threshold (2sigma, 0.05 p-value, 95% alpha threshold)
4. Find a measurable quantity which under the Null has a known distribution:
   a. If a quantity follows a known distribution, once I measure its value I can know what the probability of getting that value actually is.
   b. Different statistics:
      i. X2 statistics: difference between expectation and reality squared
      ii. Z statistics: difference between means
      iii. K-S statistics: maximum distance of cumulative distributions
5. Calculate the pivotal quantities
6. Calculate the probability of value obtained for the pivotal quantity under the Null

If probability < p-value: reject Null

Common tests and pivotal quantities



Type of question:
Do data match an expected ratio?
Is there an association between two variables?
Do samples come from the same or different populations?
Type of data: discrete, categorical (counts,frequencies)
Type of data: non-parametric (nominal, or-dinal, interval) parametric (interval)
Type of data: non-parametric parametric
a priori expectation? yes no
one variable, one sample # categories: two more than two
one variable, two or more samples, two or more categories
χ²-goodness of fit test with Yates's corr.; Fischer's exact
χ²-goodness of fit test
χ²-test for heterogeneity
two variables expl. var. under control or with smaller error than resp. var. neither one under control and with similar error
Pearson's product moment correlation
simple linear regression (**)
principal axis regression (**)
treatment variables: one two
categories two two or more
two-way non-parametric anova
data: unpaired paired
one-way anova, Kruskal-Wallis
observ. per sample less than 30 at least 30
z-test for paired data (*)
treatment variables: one two or more
categories two two or more
multi-way para-metric anova (possibly nested) (*), (**)
data: unpaired paired
one-way parametric anova (*), (**)
variances homogeneous yes no
Welch t-test
observ. per sample less than 30 at least 30
two variables, both variables have two categories at least one variable has more than two categories
χ²-test for association with Yates's corr.; Fischer's exact
χ²-test for association
two variables, data can be ranked (ordinal, interval)
Spearman's rank correlation
observ. per sample less than 30 at least 30 (and data continuous)
Mann-Whitney U test (sometimes called Wilcoxon test) (***)
z-test for unpaired data (*)
Wilcoxon's rank paired test (***)
observ. per sample less than 30 at least 30
t-test for unpaired data (*)
z-test for unpaired data (*)
t-test for paired data (*)
z-test for paired data (*)
Always check the literature for the details!
(*): requires homogeneous variances (F test)
(**): check for normally distributed residuals
(***): requires similarly shaped distributions
If you have data for which no test seems available, try to transform your data.
© Tom de Jong Frans Jacobs

Z-test

Is the mean of a sample with known variance the same as that of a known population?

Pivotal quantity

$$Z = \frac{(\bar{X} - \mu_0)}{s}$$

Sample variance

Sample mean

Population mean

$$Z \sim N(\mu=0, \sigma=1)$$

Z-test

Is the mean of a sample with known variance the same as that of a known population?

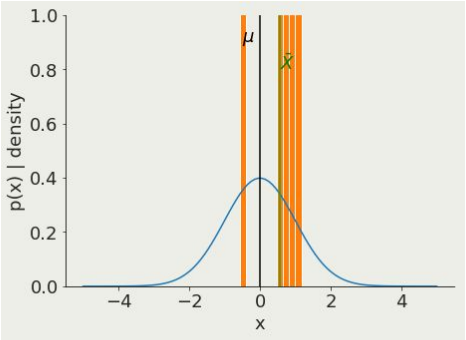
Pivotal quantity

$$Z = \frac{(\bar{X} - \mu_0)}{s}$$

Sample variance

Sample mean

Population mean



why do we need a test? why not just measuring the means and seeing it they are the same?

$$Z \sim N(\mu=0, \sigma=1)$$

Z-test

Is the mean of a sample with known variance the same as that of a known population?

Pivotal quantity

$$Z = \frac{(\bar{X} - \mu_0)}{s}$$

Sample variance

Sample mean

Population mean

$$Z \sim N(\mu=0, \sigma=1)$$

The Z test provides a trivial interpretation of the measured quantity: the Z value is exactly the distance for the mean of the standard distribution of possible outcomes in units of standard deviation

so a result of 0.13 means we are 0.13 standard deviations to the mean (p>0.05)

t-test

Are the means of 2 samples significantly different?

Pivotal quantity

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{\Delta}}}$$

t ~ Student's t $\left(\mathrm{d.f.} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1} + \frac{\left(s_2^2/n_2\right)^2}{n_2-1}} \right)$

unbiased variance estimator

$$s_{\bar{\Delta}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

size of sample

t-test

Are the means of 2 samples significantly different?

Pivotal quantity

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{\Delta}}}$$

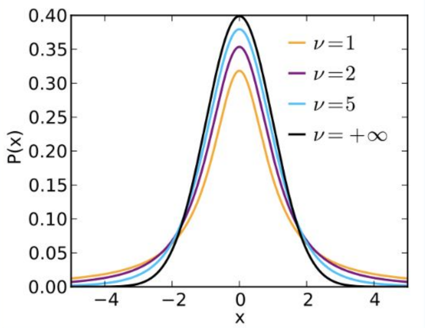
$$s_{\bar{\Delta}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

unbias variance estimator

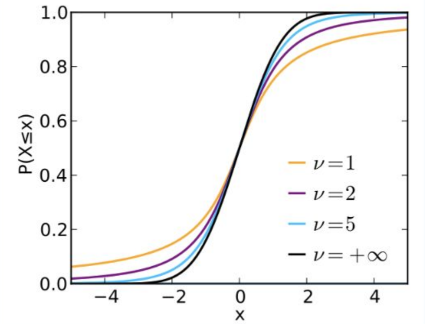
size of sample

Student's *t*

Probability density function



Cumulative distribution function



Parameters	$\nu > 0$ degrees of freedom (real)
Support	$x \in (-\infty, \infty)$
PDF	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$
CDF	$\frac{1}{2} + x\Gamma\left(\frac{\nu+1}{2}\right) \times \frac{{}_2F_1\left(\frac{1}{2},\frac{\nu+1}{2};\frac{3}{2};-\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}$ where ${}_2F_1$ is the hypergeometric function
Mean	0 for $\nu > 1$, otherwise undefined

Parameters	$\nu > 0$ degrees of freedom (real)
Support	$x \in (-\infty, \infty)$
PDF	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$
CDF	$\frac{1}{2} + x\Gamma\left(\frac{\nu+1}{2}\right) \times \frac{{}_2F_1\left(\frac{1}{2},\frac{\nu+1}{2};\frac{3}{2};-\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}$ where ${}_2F_1$ is the hypergeometric function
Mean	0 for $\nu > 1$, otherwise undefined
Median	0
Mode	0
Variance	$\frac{\nu}{\nu-2}$ for $\nu > 2$, ∞ for $1 < \nu \le 2$, otherwise undefined

t-test

To interpret the outcome of a t-test I have to figure out the probability of a given p

Are the means of 2 samples significantly different?

Pivotal quantity

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{\Delta}}}$$

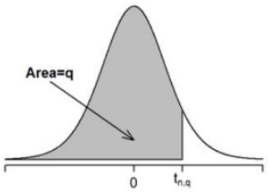
unbias variance estimator

$$s_{\bar{\Delta}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

size sample

Quartiles of the t Distribution

The table gives the value if $t_{n;q}$ - the qth quantile of the t distribution for n degrees of freedom



	q = 0.6	0.75	0.9	0.95	0.975	0.99	0.995	0.9975	0.999	0.9995
n = 1	0.3249	1.0000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.2887	0.8165	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.2767	0.7649	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.2707	0.7407	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.2672	0.7267	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.2648	0.7176	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.2632	0.7111	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.2619	0.7064	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.2610	0.7027	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.2602	0.6998	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.2596	0.6974	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.2590	0.6955	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.2586	0.6938	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.2582	0.6924	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140

χ2 test

are the data what is expected from the model - there are a few χ2 tests. The one here is the "Pearson's χ2 tests"

Pivotal quantity

$$X^2 \equiv \sum_i \frac{(f(x_i)-y_i)^2}{\sigma_i^2}$$

model

observation

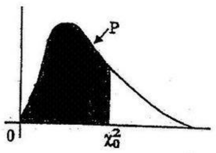
uncertainty

N of observations

$$\chi^2 \sim \chi^2(df = n-1)$$

N parameters model

χ^2 Distribution



The table below gives the value x_0^2 for which $P[x^2 < x_0^2] = P$ for a given number of degrees of freedom and a given value of P.

Degrees of Freedom	Values of P									
	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.01	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997

Kolmogorov–Smirnov test

Kolmogorov-Smirnov test:

do two samples come from the same parent distribution?

Pivotal quantity

$$D_n = \sup_x |F_n(x) - F(x)|$$

Cumulative distribution 1

Cumulative distribution 2