

Statistics & Statistical Machine Learning

Bài 2: Inferential Statistics



Quang-Khai Tran, Ph.D
CyberLab, 04/2022



(Ảnh: Internet)

Nội dung



1. Tổng Quan
2. Hypothesis Testing
3. Statistical Tests
4. A/B Testing



Phần 1. Tổng quan

- 1.1. Giới thiệu Thống kê Suy luận
- 1.2. So sánh với Thống kê Mô tả
- 1.3. Ba vấn đề chính trong TKSL

1.1 Giới thiệu Thống kê Suy luận

❖ Descriptive Statistics & Inferential Statistics

Descriptive Statistics

(là khoa học về việc) thu thập, sắp xếp, tóm tắt và hiển thị dữ liệu

Inferential Statistics

(là khoa học về việc) sử dụng descriptive statistics để ước lượng (estimate) các tham số của quần thể

1.1 Giới thiệu Thống kê Suy luận

❖ 2 loại phân tích (analysis)

Exploratory Analysis (phân tích khai phá)
sử dụng data để ước lượng (estimate) các tham số

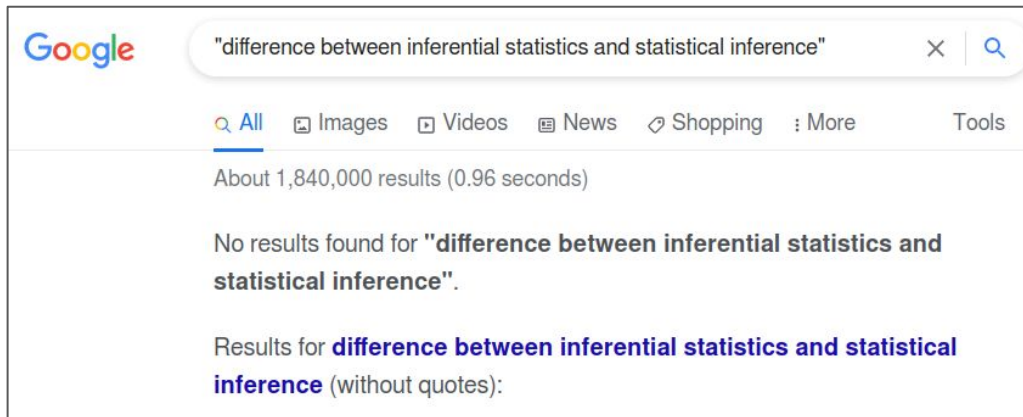
Confirmatory Analysis (phân tích khẳng định)
sử dụng thống kê để kiểm định các phát biểu (hoặc các khẳng định, các nhận định) về thế giới thực (hypothesis/hypotheses)

1.1 Giới thiệu Thống kê Suy luận

Các bước thực hiện một nghiên cứu (thử nghiệm) thống kê:
(thuật ngữ: statistical study/experimental study, experiment)

1. Xác định các hạng mục cơ bản (design):
 - Phát biểu câu hỏi cần nghiên cứu
 - Xác định population và các variables
 - Xác định phương pháp lấy mẫu
2. Thu thập dữ liệu
3. Sắp xếp, tổ chức dữ liệu
4. Phân tích dữ liệu để trả lời câu hỏi trên

❖ Statistical Inference vs. Inferential Statistics



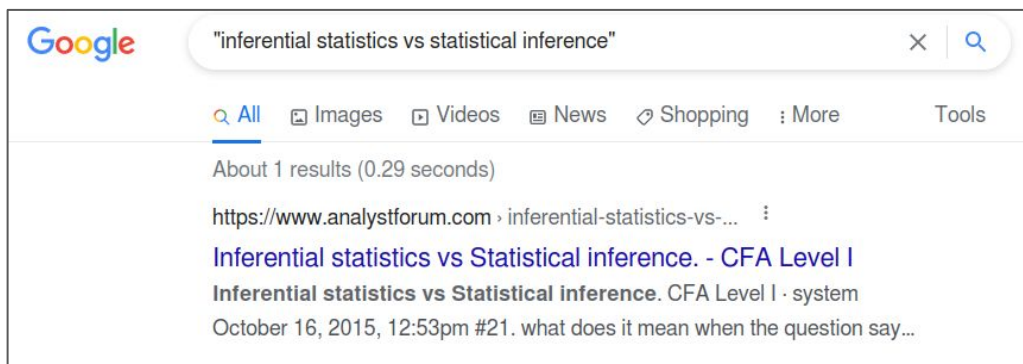
Google "difference between inferential statistics and statistical inference" X Q

All Images Videos News Shopping : More Tools

About 1,840,000 results (0.96 seconds)

No results found for **"difference between inferential statistics and statistical inference"**.

Results for **difference between inferential statistics and statistical inference** (without quotes):



Google "inferential statistics vs statistical inference" X Q

All Images Videos News Shopping : More Tools

About 1 results (0.29 seconds)

<https://www.analystforum.com> > inferential-statistics-vs-... :
Inferential statistics vs Statistical inference. - CFA Level I
Inferential statistics vs Statistical inference. CFA Level I · system
October 16, 2015, 12:53pm #21. what does it mean when the question say...

❖ Statistical Inference vs. Inferential Statistics

"Statistical inference is the process of using data analysis to infer properties of an underlying distribution of probability"

(là quá trình sử dụng phân tích dữ liệu để suy ra các thuộc tính của một phân bố xác suất)

"Inferential statistical analysis infers properties of a population, for example by testing hypotheses and deriving estimates. It is assumed that the observed data set is sampled from a larger population"

(là dạng phân tích) suy luận ra thuộc tính của một quần thể, ví dụ bằng cách kiểm tra các giả thuyết và tính toán ra các ước lượng

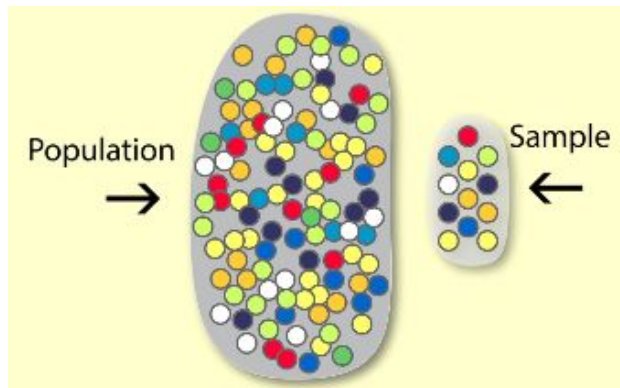
Source: https://en.wikipedia.org/wiki/Statistical_inference

❖ Statistical Inference vs. Inferential Statistics

- Statistical Inference: (Sự) Suy luận (có tính) thống kê
 - Là một hoạt động (hành động) (hoặc quá trình - process)
 - Quá trình này nhằm đưa ra **kết luận** (conclusion) / **tuyên bố** (proposition) về một quần thể dựa trên các phương pháp phân tích thống kê
- Inferential Statistics: Thống kê (có tính) suy luận
 - Gồm các phương pháp/kỹ thuật (để thực hiện hành động/quá trình ở trên)
 - Bao gồm các phương pháp thống kê

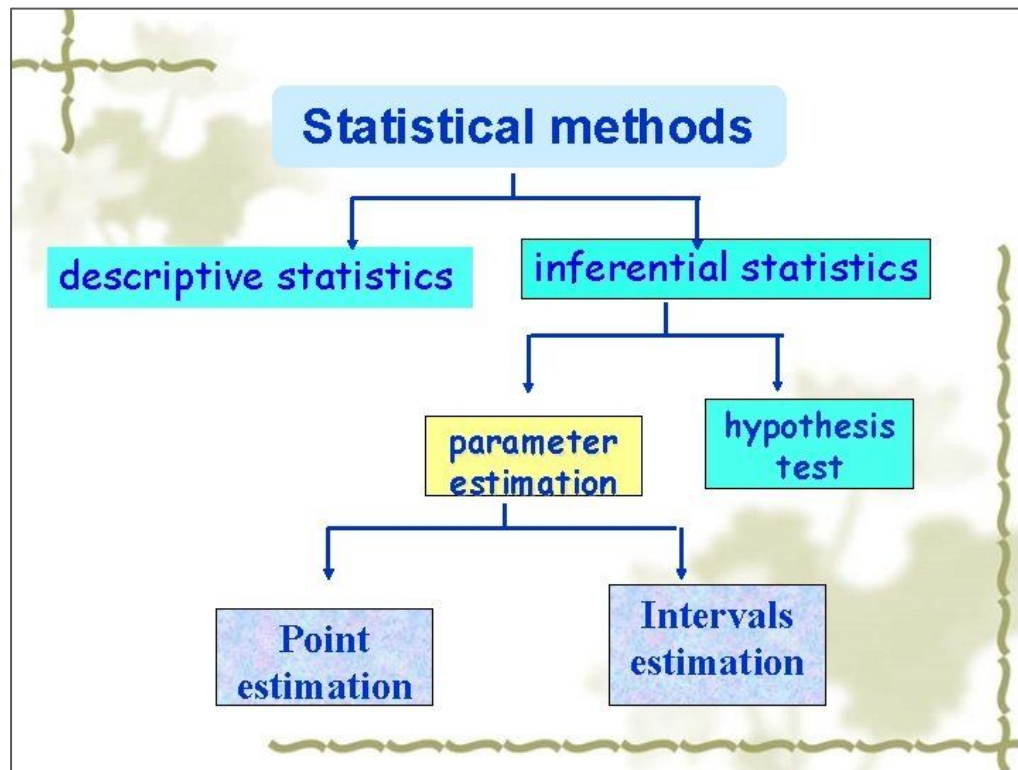
1.1 Giới thiệu Thống kê Suy luận

- ❖ Descriptive statistics chỉ quan tâm tới việc đưa ra kết luận về dữ liệu hiện có
- ❖ Inferential Statistics:
 - Giúp đưa ra kết luận xa hơn so với dữ liệu hiện có
⇒ Về toàn bộ quần thể (tức model đã sinh ra dữ liệu)
 - Giúp đưa ra dự đoán (prediction, inference). Trong Machine Learning:
 - Training/learning: suy ra (tìm ra) các đặc tính của model
 - Inference/prediction: sử dụng model để tạo ra dự đoán



1.1 Giới thiệu Thống kê Suy luận

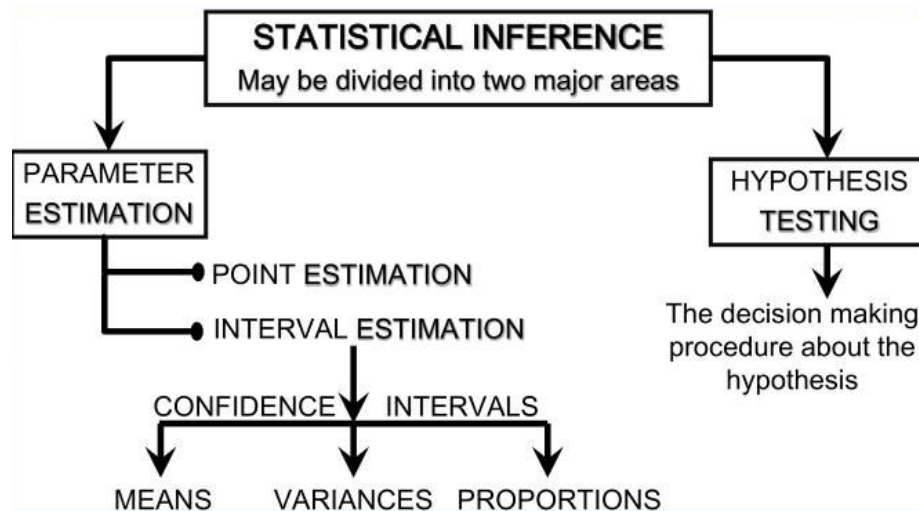
- ❖ Hai cách tiếp cận chính của Inferential Statistics



1.1 Giới thiệu Thống kê Suy luận

Hai cách tiếp cận chính

- ❖ Parameter Estimation
(Ước lượng/tính tham số)
 - Suy đoán ra giá trị của tổng thể bằng cách lấy mẫu
- ❖ Hypothesis Testing
(Kiểm định giả thuyết thống kê)
 - Ra quyết định dựa trên giả thuyết



Tham khảo về tổng quan 02 cách tiếp cận này (tiếng Việt):

https://rpubs.com/tranquangquy_ictu/783476

1.1 Giới thiệu Thống kê Suy luận

- ❖ Hai cách tiếp cận chính theo Frequentist vs. Bayesian

	Frequentist	Bayesian
Hypothesis test	p value (null hypothesis significance test)	Bayes factor
Estimation with uncertainty	maximum likelihood estimate with confidence interval (The "New Statistics")	posterior distribution with highest density interval

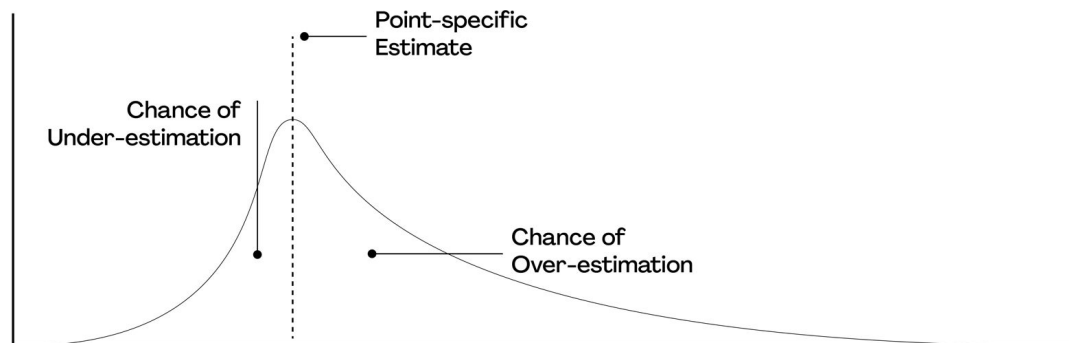
1.1 Giới thiệu Thống kê Suy luận

Parameter Estimation (Ước lượng/tính tham số)

- ❖ Dựa vào khái niệm tần số (frequency) và lấy mẫu (sampling)
- ❖ Chia nhỏ làm 02 loại:
 - Ước lượng điểm (point estimate): một con số độc lập (single point) được xem là tốt nhất hay gần với giá trị cần dự đoán nhất
 - Ước lượng khoảng (interval estimate): một khoảng có chặn trên và chặn dưới mà người ta tin rằng giá trị thực sự cần ước lượng sẽ nằm trong khoảng đó

Tham khảo (tiếng Việt):

<https://vimentor.com/vi/blog/khoang-tin-cay-confidence-interval>

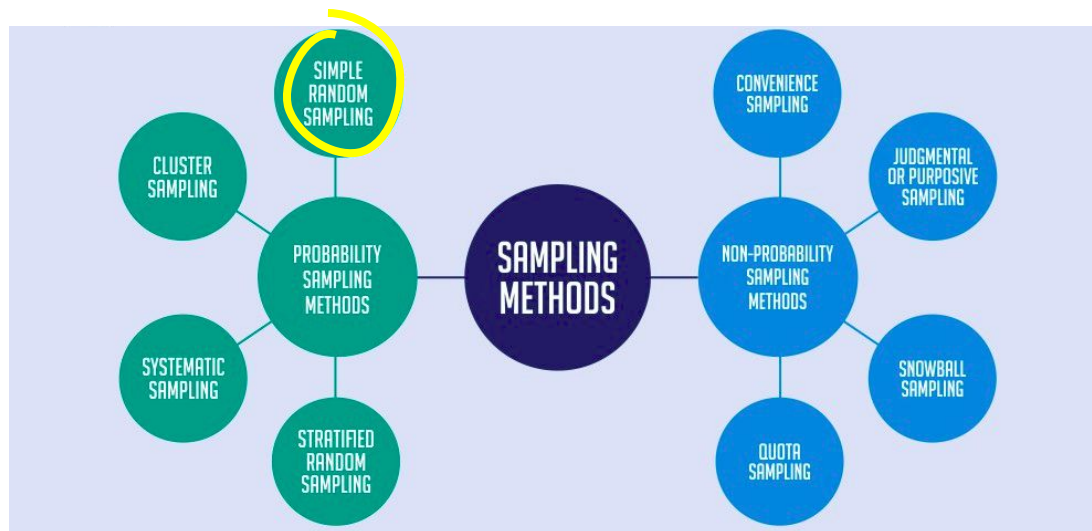


1.2 So sánh với Thống kê Mô tả

	Thống kê mô tả	Thống kê suy luận
Ý nghĩa	Là nhánh thống kê liên quan đến việc mô tả mẫu đang nghiên cứu	Tập trung vào việc rút ra kết luận về tổng thể, dựa trên cơ sở phân tích và quan sát mẫu
Thực hiện	Tổ chức, phân tích và trình bày dữ liệu một cách có ý nghĩa	So sánh, kiểm tra và dự đoán dữ liệu
Kết quả	Biểu đồ, đồ thị và bảng	Xác suất và quyết định
Sử dụng	Để mô tả một tình huống	Để giải thích các cơ hội xảy ra của một sự kiện
Chức năng	Giúp giải thích dữ liệu đã được biết, để tóm tắt mẫu	Giúp đưa ra kết luận để tìm hiểu về tổng thể, vượt ra ngoài dữ liệu có sẵn

1.3 Ba vấn đề chính trong TKSL

1. Một sample có vẻ sẽ là **một đại diện tốt** của population
2. Tuy nhiên, sẽ luôn có một **thành phần "không chắc chắn"** (uncertainty)
 - Ẩn chứa trong dữ liệu và sample
 - Thành phần này sẽ ảnh hưởng đến mức độ đại diện đó
3. Do đó, **việc lấy mẫu** (sampling) cũng là điều cần quan tâm





Phần 2. Hypothesis Testing

2.1. Giới thiệu Hypothesis Testing

2.2. Hai loại giả thuyết:

- Null hypothesis
- Alternative hypothesis

2.3. Two-tailed Test vs. One-tailed Test

2.4. Các giá trị cần xác định

(Significance value, Critical value, p-Value, test-statistic)

2.5. Error Type I & II

2.1 Giới thiệu Hypothesis Testing

Statistical Hypothesis Testing (Kiểm định giả thuyết thống kê)

- ❖ Là một thủ tục chuẩn tắc (formal procedure)/hoặc phương pháp
- ❖ Dùng để kiểm chứng tính đúng đắn của:
 - một phát biểu (claim), hoặc
 - sự tin tưởng (belief)
 - về tham số của quần thể/tổng thể (population parameter)

Lưu ý: "Giả thiết" != "Giả thuyết"

Giả thiết (assumption): điều cho trước để làm căn cứ suy luận

Giả thuyết (hypothesis): là luận điểm, nhận định cần được chứng minh

2.1 Giới thiệu Hypothesis Testing

Hai nhóm chủ thể trong nghiên cứu thống kê:

- ❖ Control group (nhóm kiểm soát): là nhóm các đối tượng (hoặc chủ thể, subject) không nhận được thay đổi gì mới trong thí nghiệm
- ❖ Treatment/experimental group (nhóm thí nghiệm): là nhóm các đối tượng nhận được các thay đổi mà nhà nghiên cứu muốn kiểm tra

Biến số gây nhiễu (*confounding variables*)

- Là các yếu tố không đo đếm được trong thí nghiệm
- Các yếu tố này cũng có ảnh hưởng đến các chủ thể (ngoài những thay đổi, treatment, có chủ đích ở trên)

5 bước cơ bản để thực hiện Hypothesis Testing:

1. Phát biểu các giả thuyết:
 - Null hypothesis (H_0)
 - Alternative hypothesis (H_1)
2. Thu thập dữ liệu (được thiết kế riêng cho việc kiểm định)
 - Xem xét giả thiết (assumptions, nếu cần)
 - Ví dụ: giả thiết về tính độc lập thống kê, hay dạng phân bố của dữ liệu
3. Thực hiện một kiểm định thống kê phù hợp
4. Quyết định "Loại trừ" hoặc "Không thể loại trừ" giả thuyết H_0
5. Thể hiện kết quả và thảo luận điều đã rút ra

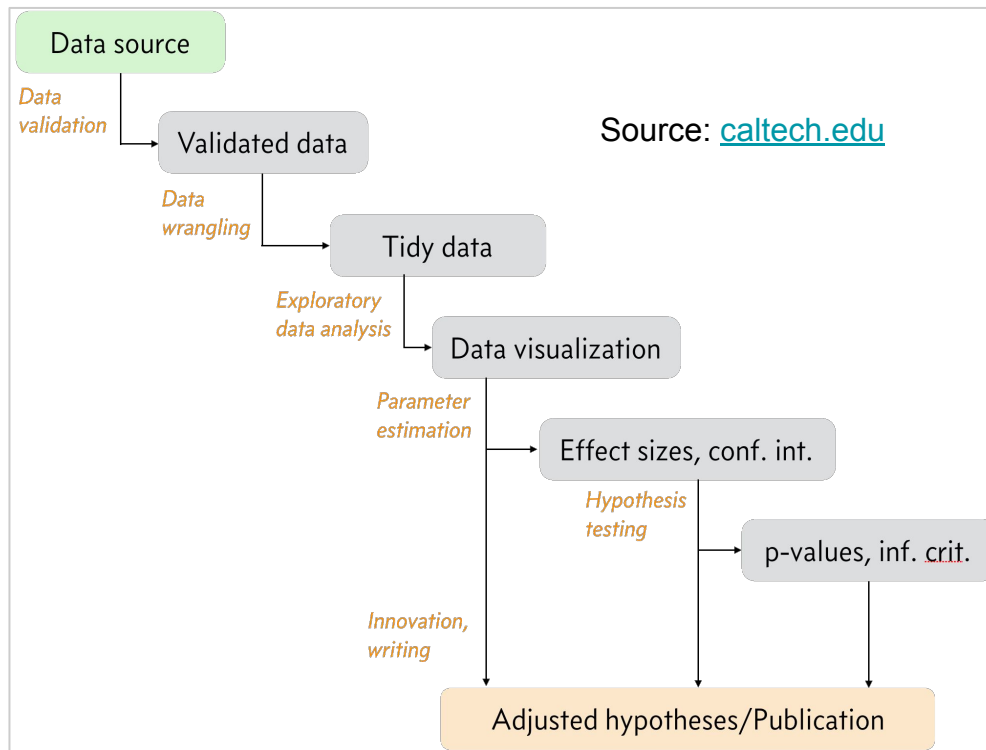
Bước 3: Thực hiện statistical test. Có các loại chính sau sau

Phân loại	Các loại thành phần
Parametric Test	Z Test / T Test / F Test / ANOVA
Non-parametric Test	Chi-square Test / Wilcoxon
Khác	A/B Testing

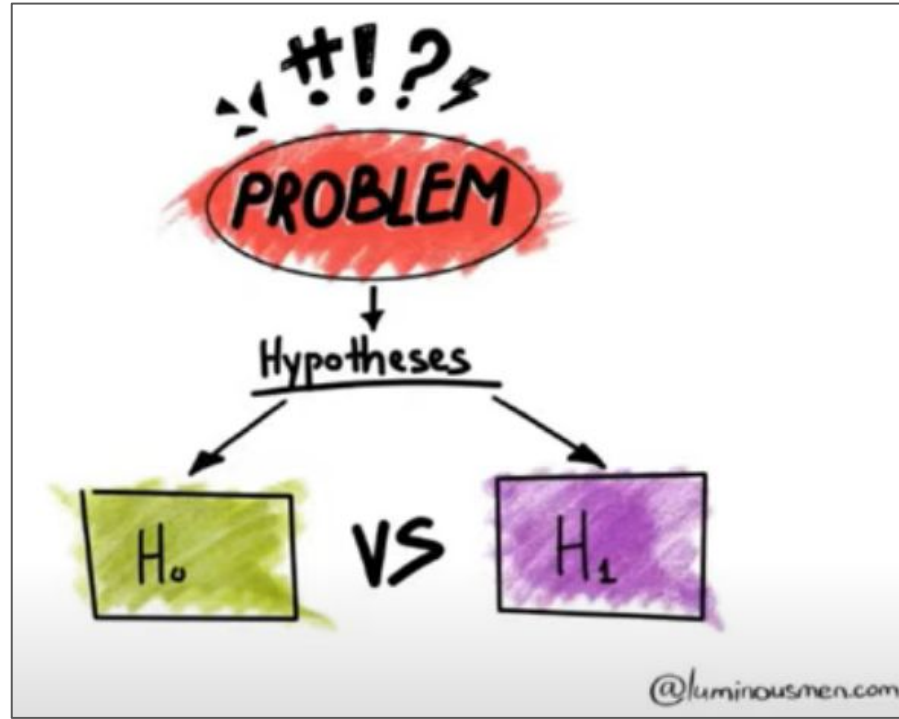
Link tham khảo:

<https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/>
https://en.wikipedia.org/wiki/Test_statistic

❖ Data Analysis Pipeline with Hypothesis Testing



2.2 Hai loại giả thuyết



- ❖ Giả thuyết không (null hypothesis, H_0)
 - Giả định rằng trong một tập hợp các quan sát không có ý nghĩa thống kê nào
 - Hay nói cách khác: không có sự khác biệt giữa 2 tập dữ liệu được quan sát
 - H_0 được (mặc định) cho là đúng cho đến khi có bằng chứng thống kê bác bỏ nó

Kết luận của một Hypothesis Testing trên H_0 :

1. Reject H_0 (loại bỏ/khước từ H_0)
2. Fail to reject H_0 (không đủ cơ sở để loại bỏ H_0)

Tham khảo: Tài liệu tiếng Việt

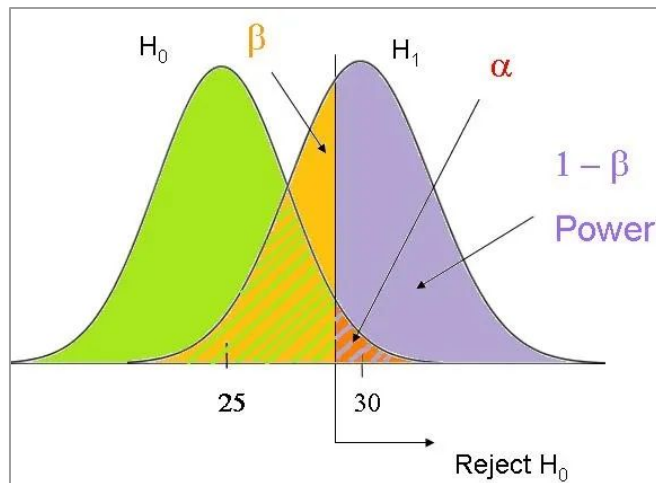
<https://vietnambiz.vn/gia-thuyet-khong-null-hypothesis-la-gi-vi-du-ve-gia-thuyet-khong-20191113092744946.htm>

2.2 Hai loại giả thuyết

- ❖ Giả thuyết thay thế (alternative hypothesis: H_A , H_1)
 - Đại diện cho một điều trái ngược (phản bác) lại niềm tin về tham số tổng thể
 - Thường là phát biểu ngược lại với H_0

Tham khảo: Tài liệu tiếng Việt

<https://vietnambiz.vn/gia-thuyet-khong-null-hypothesis-la-gi-vi-du-ve-gia-thuyet-khong-20191113092744946.htm>

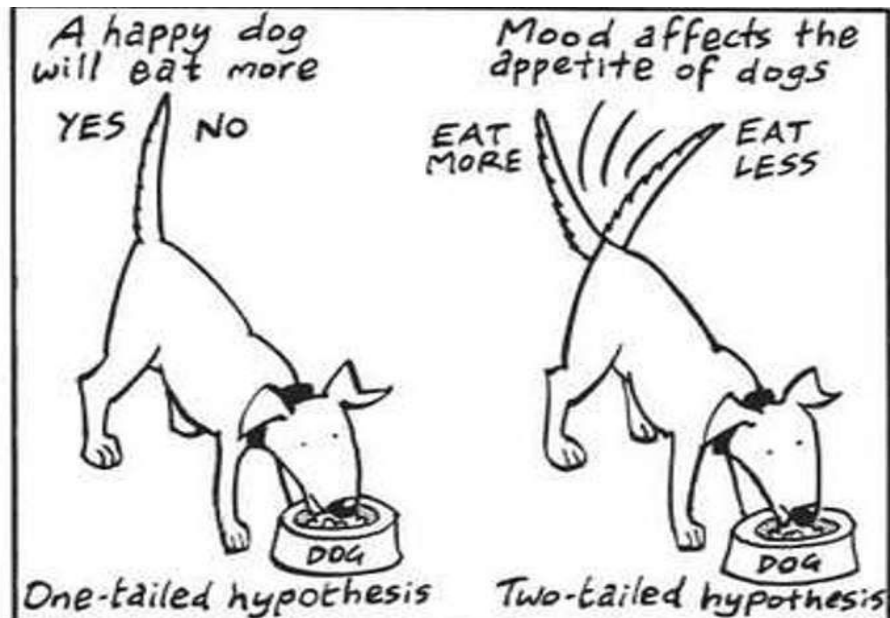


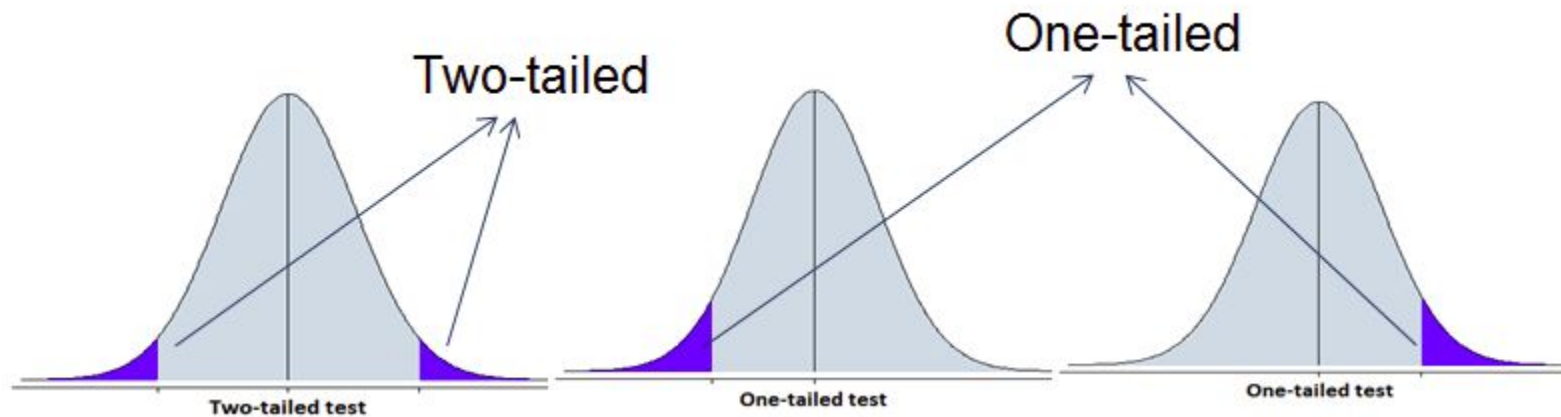
2.2 Hai loại giả thuyết

Warning: Coi chừng hiểu sai!



Link: <https://study.com/academy/lesson/effect-size-in-hypothesis-testing-definition-interpretation.html>

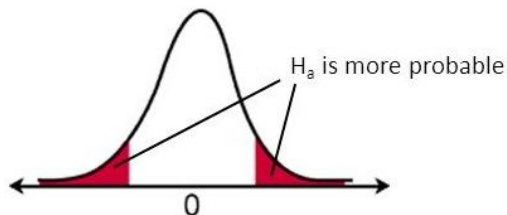
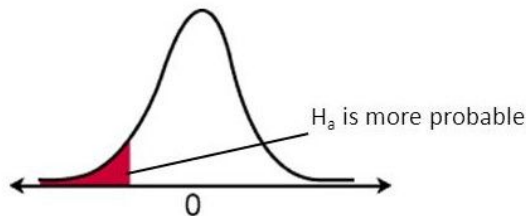
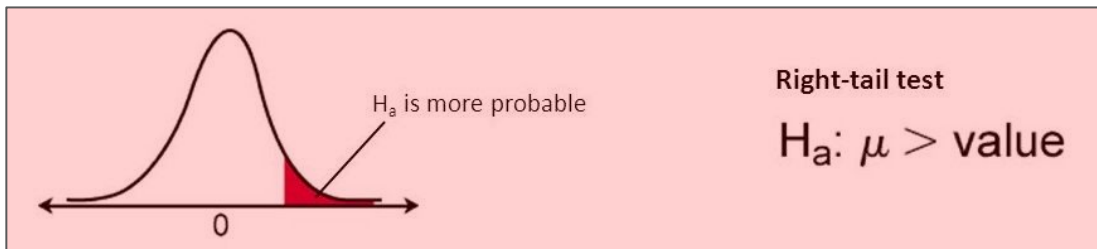




2.3 Two-tailed Test vs. One-tailed Test

One-tailed (right):

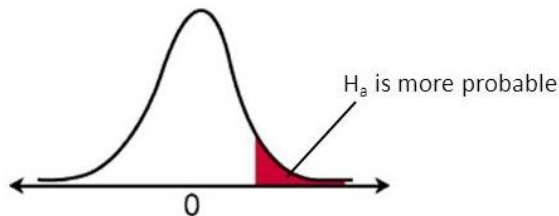
1. $H_0 \leq \text{value of parameter}$
2. $H_1 > \text{value of parameter}$



2.3 Two-tailed Test vs. One-tailed Test

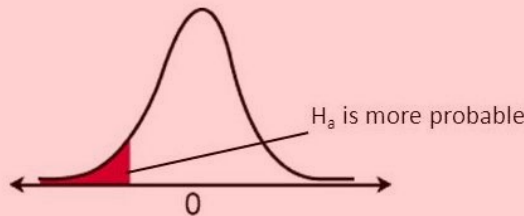
One-tailed (left):

1. $H_0 \geq \text{value of parameter}$
2. $H_1 < \text{value of parameter}$



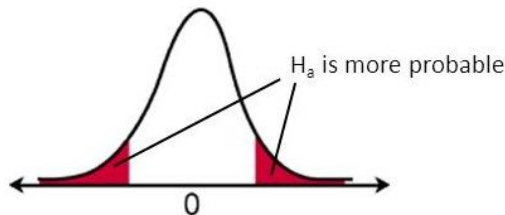
Right-tail test

$$H_a: \mu > \text{value}$$



Left-tail test

$$H_a: \mu < \text{value}$$



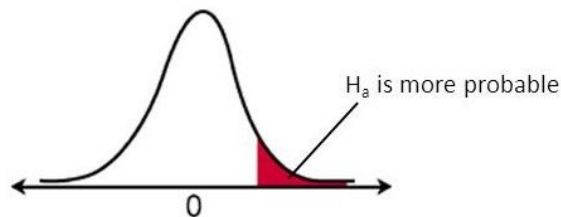
Two-tail test

$$H_a: \mu \neq \text{value}$$

2.3 Two-tail Test vs. One-tail Test

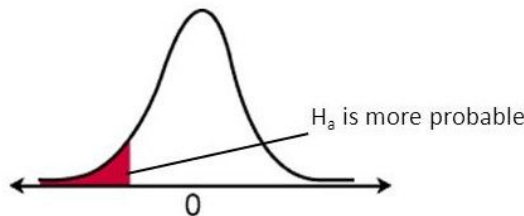
Two-tailed:

1. H_0 = value of parameter
2. $H_1 \neq$ value of parameter



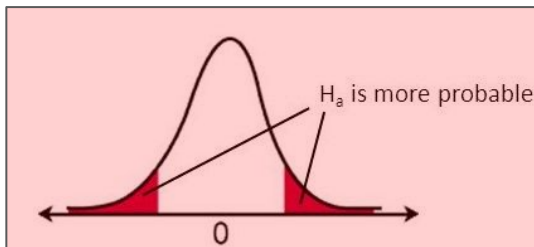
Right-tail test

$$H_a: \mu > \text{value}$$



Left-tail test

$$H_a: \mu < \text{value}$$

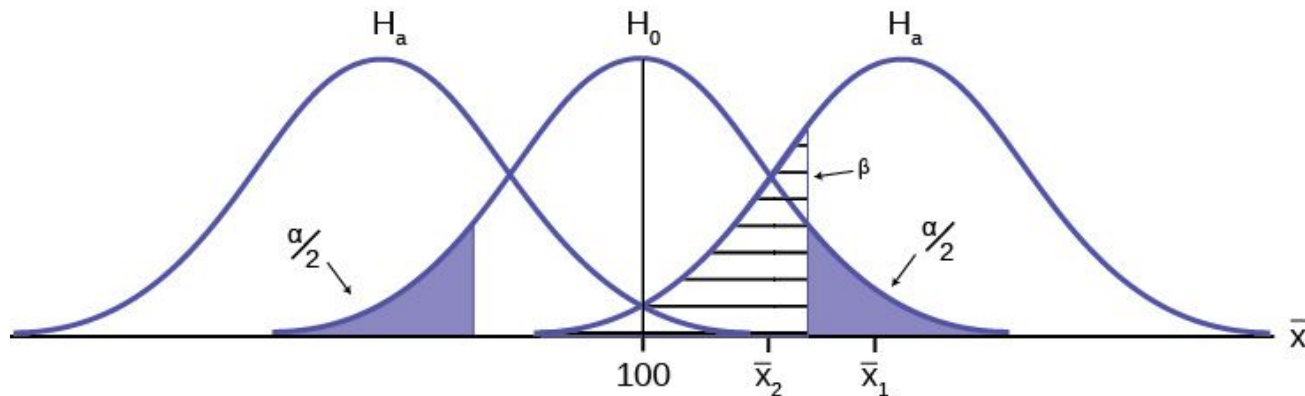


Two-tail test

$$H_a: \mu \neq \text{value}$$

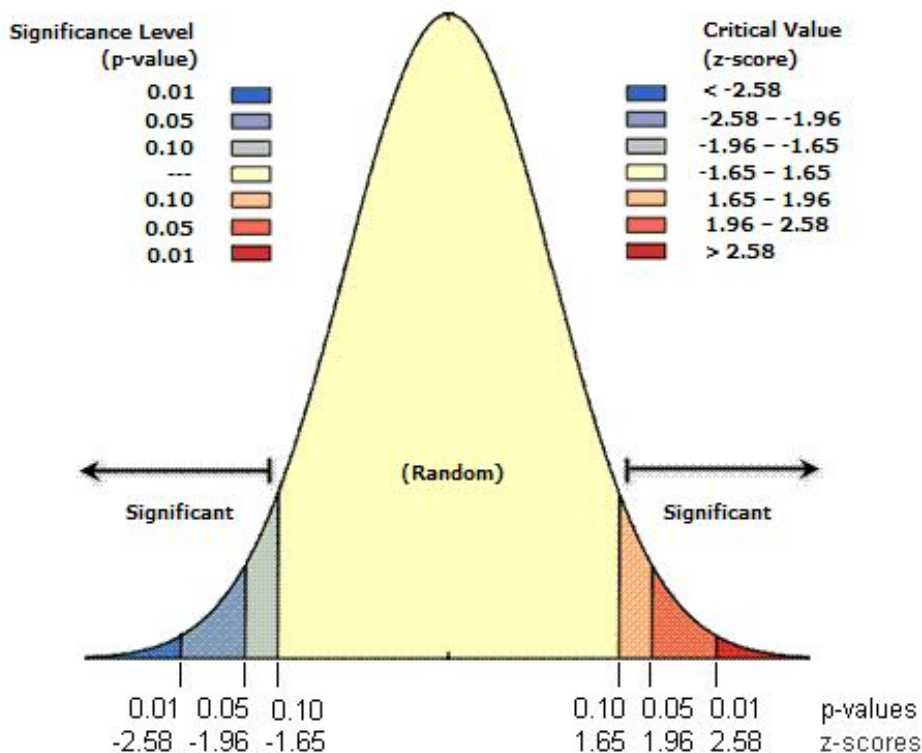
2.3 Two-tailed Test vs. One-tailed Test

❖ Two-tailed



$$H_0: \mu = 100$$
$$H_a: \mu \neq 100$$

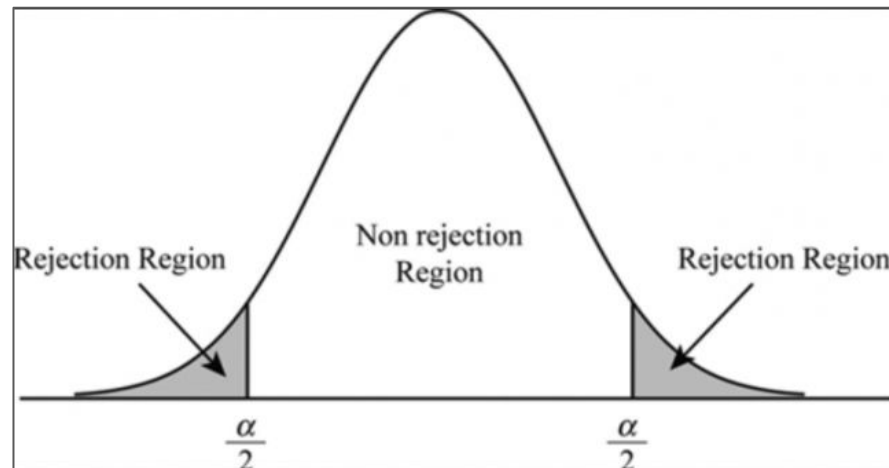
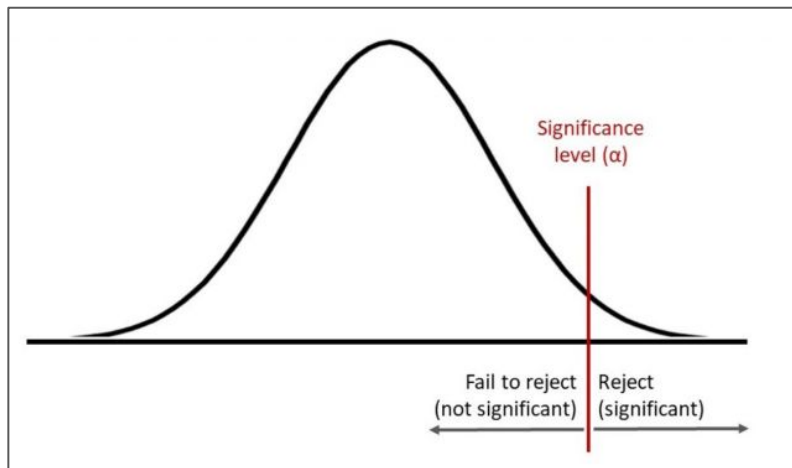
- ❖ Sinificance Level
- ❖ Critical value
- ❖ Test statistic
- ❖ p-Value



2.4 Các giá trị cần xác định

❖ Giá trị Significance Level (Mức ý nghĩa của thống kê)

- Một kết luận cho rằng kết quả từ kiểm định hoặc thử nghiệm không xảy ra do ngẫu nhiên hay tình cờ, thay vào đó là do một nguyên nhân cụ thể
- Thường được ký hiệu là α
(giá trị mà ta định trước: 0.01, 0.05, 0.1, tương ứng 1%, 5%, 10%)
- Sẽ dùng để so sánh với p-Value

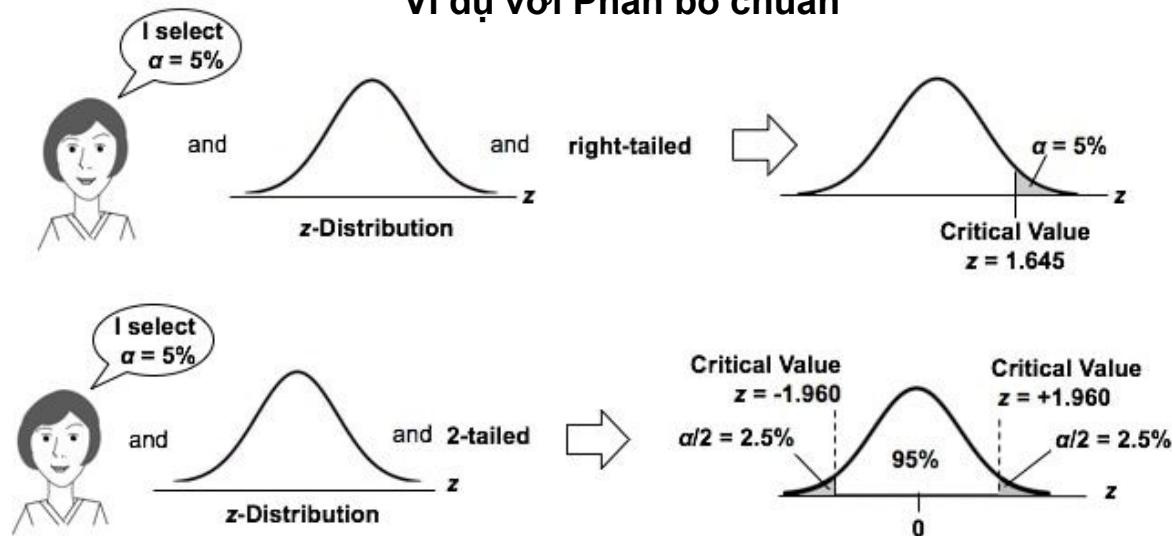


2.4 Các giá trị cần xác định

❖ Giá trị Critical value

- Là giá trị được suy ra từ Significance Level
- Tùy theo loại kiểm định, giá trị là: z-score, t, F, Chi-square

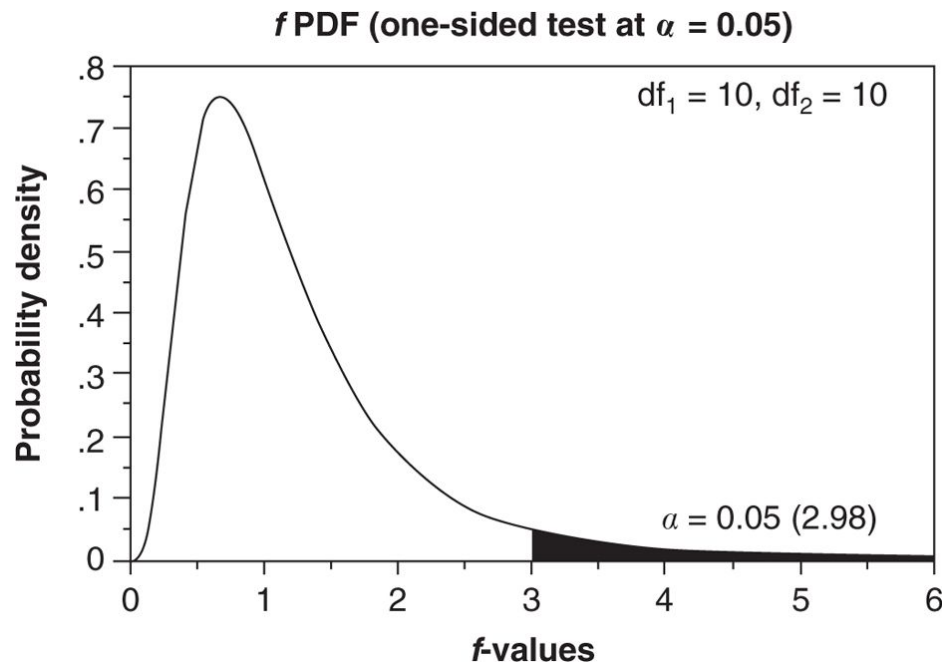
Ví dụ với Phân bố chuẩn



2.4 Các giá trị cần xác định

❖ Giá trị Critical value

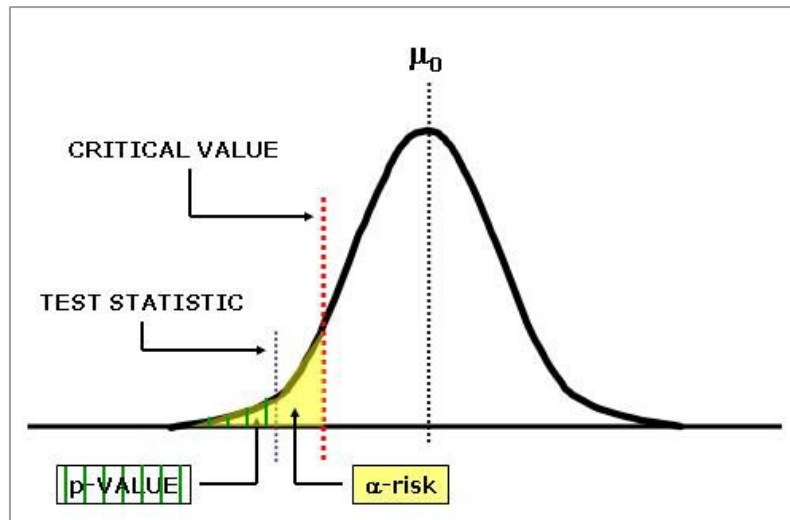
- Ví dụ: với một F-distribution như hình dưới
 - Khi $\alpha = 0.05$ thì $CV = 2.98$ (f-score = 2.98)



2.4 Các giá trị cần xác định

❖ Giá trị Test statistic

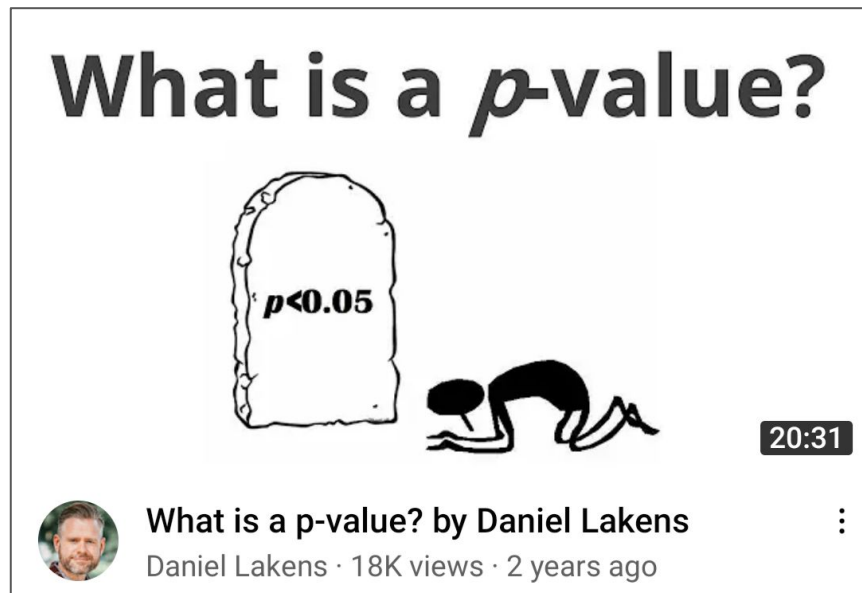
- Là giá trị ta cần kiểm tra (so sánh với critical value)
- Tùy theo loại kiểm định, giá trị có thể là: z-score, t, F, Chi-square
- Thường dùng để tính p-Value



2.4 Các giá trị cần xác định

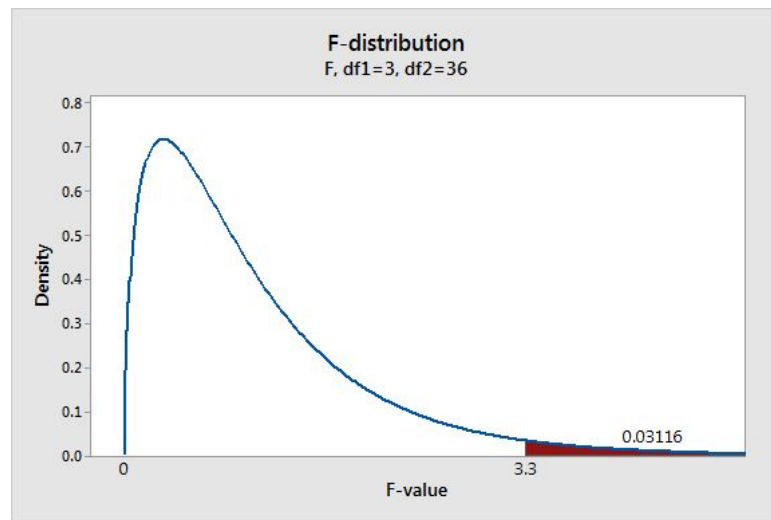
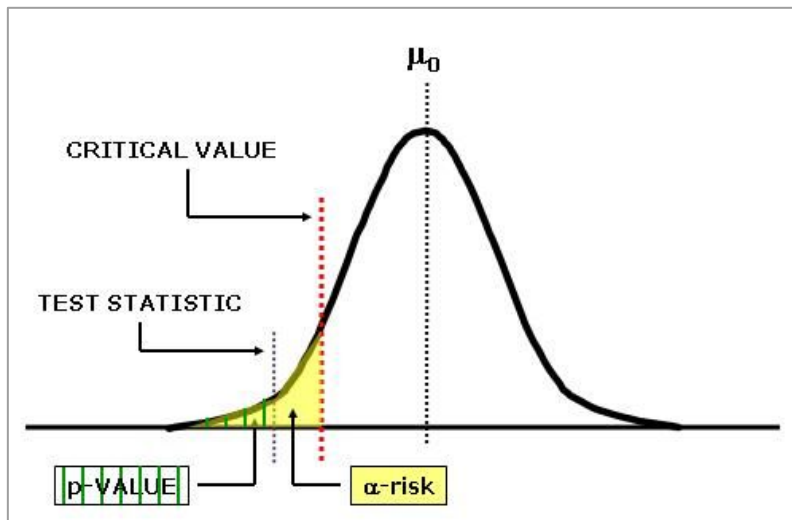
❖ Giá trị p-value

- p-value là xác suất dữ liệu xảy ra trong điều kiện giả thuyết H_0 là đúng
- Do đó nếu p-value thấp thì chúng ta sẽ bác bỏ giả thuyết vô hiệu



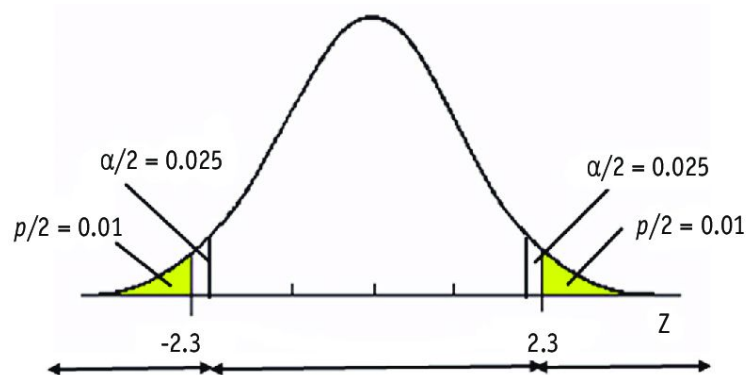
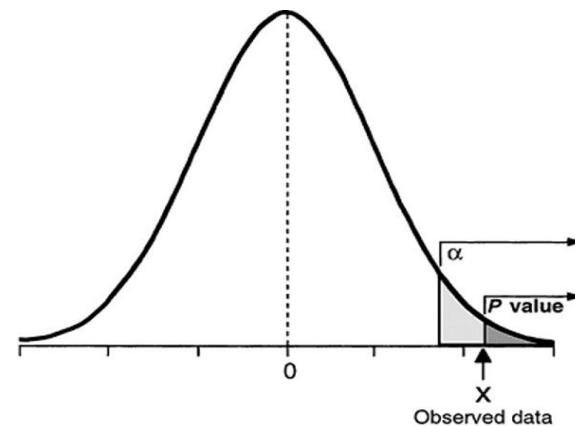
2.4 Các giá trị cần xác định

- ❖ p-value là giá trị được tạo thành bởi tổng của 3 thành phần (xác suất):
 1. Xác suất mà sự ngẫu nhiên tạo ra dữ liệu quan sát được (test statistic)
 2. Xác suất quan sát được các dữ liệu "hiếm như" (equally rare) xác suất trên
 3. Xác suất quan sát được các dữ liệu "hiếm hơn" (rarer) xác suất (1)



2.4 Các giá trị cần xác định

- ❖ p-value: kết luận từ p-Value



2.5 Type-I & Type-II errors

- ❖ Có 2 loại lỗi có thể xảy ra







	Null hypothesis is TRUE	Null hypothesis is FALSE
Reject null hypothesis	Type I Error (False positive)	Correct outcome! (True positive)
Fail to reject null hypothesis	Correct outcome! (True negative)	Type II Error (False negative)

2.5 Type-I & Type-II errors

Cách phân biệt:

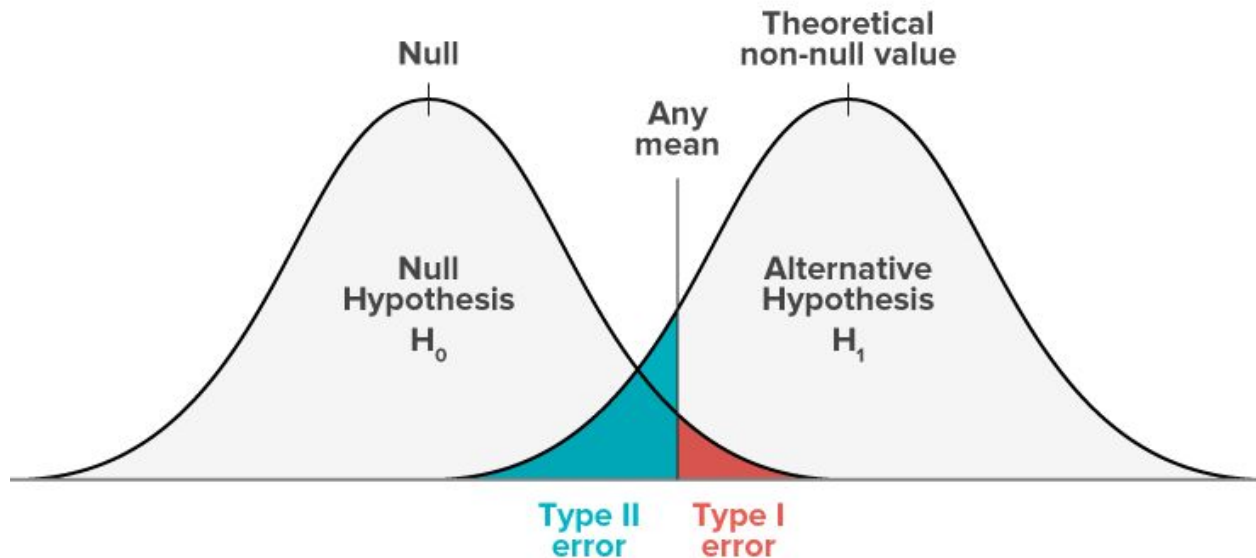
- ❖ Bị kết tội "oan": One
- ❖ Ngược lại (được tha trong khi có tội): Two

		The Person is	
		Innocent	Guilty
The Judge Says	Innocent	 No Error	 Type 2 error
	Guilty	 Type 1 error	 No Error

2.5 Type-I & Type-II errors

Cách phân biệt:

- ❖ H_0 đúng lại kết luận là sai: One
- ❖ H_0 sai lại kết luận là đúng: Two



Link: <https://rpubs.com/tmcurley/onewayanova>



Phần 3. Các loại Statistical Tests

3.1. Parametric Test

(kiểm định tham số)

- T-Test / Z-Test / F-Test
- ANOVA Test

3.2. Non-parametric Test

(Kiểm định phi tham số)

- Chi-squared Test

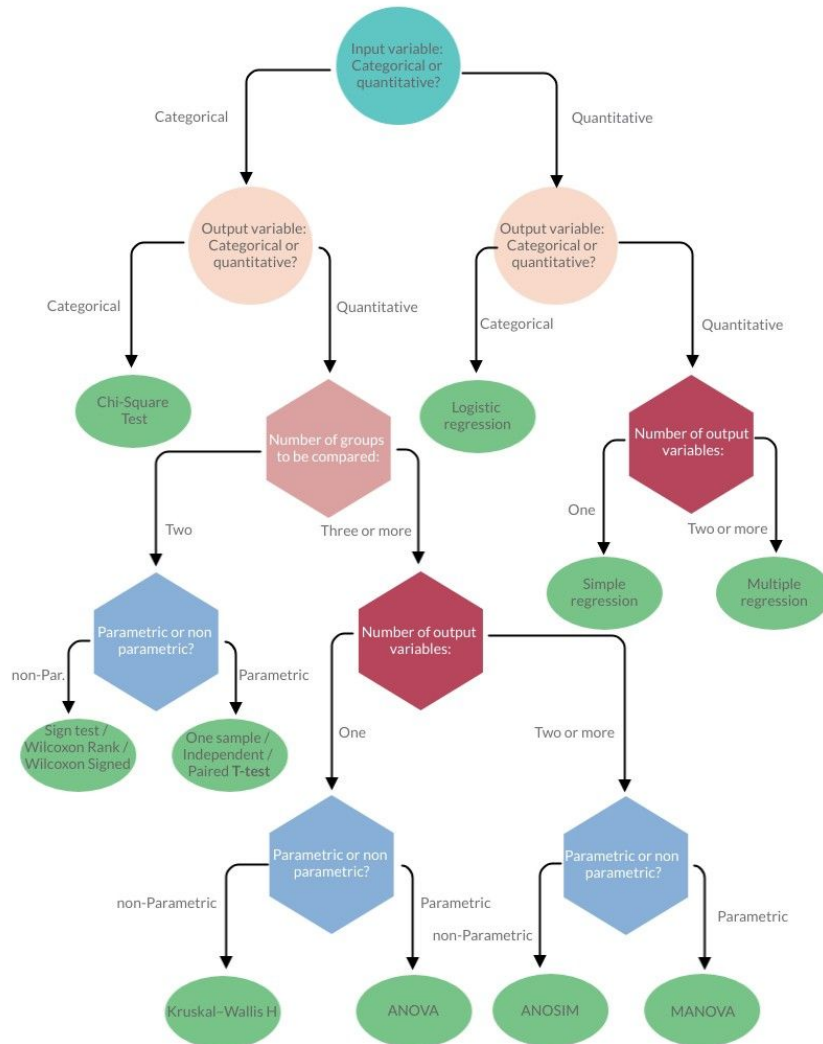
3.3. A/B Test

- ❖ Các kiểm định thống kê (statistical tests) là các kỹ thuật/phương pháp kiểm định cụ thể được sử dụng trong kiểm định giả thuyết

- ❖ Tham khảo:

<https://towardsdatascience.com/a-nova-t-test-and-other-statistical-tests-with-python-e7a36a2fdc0c>

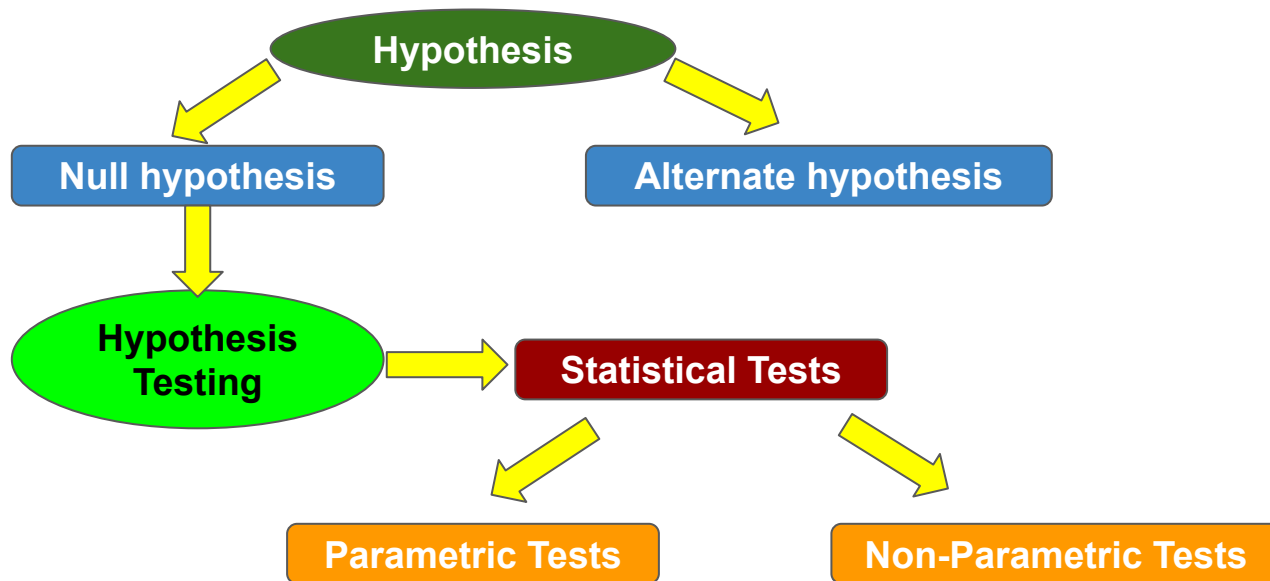
[https](https://towardsdatascience.com/a-nova-t-test-and-other-statistical-tests-with-python-e7a36a2fdc0c)



Phân loại hai loại kiểm định chính:

- ❖ Parametric Test (kiểm định có tham số):
 - Dựa trên các giả định (assumption) về phân bố của dữ liệu
 - Thông tin về phân bố của quần thể là đã biết (mean, variance là cố định)
 - Ví dụ: Z-Test, T-Test, ANOVA (và rất nhiều phương pháp khác)
- ❖ Non-parametric Test (kiểm định phi tham số):
 - Không cần dựa trên giả định về phân bố,
 - Thông tin về phân bố của quần thể là chưa biết trước
 - Các tham số (parameters) là không cố định
 - Có thể thực hiện với bất cứ phân bố nào
 - Ví dụ: Chi-square Test, Wilcoxon (và rất nhiều phương pháp khác)

Phân loại hai loại kiểm định chính (1):

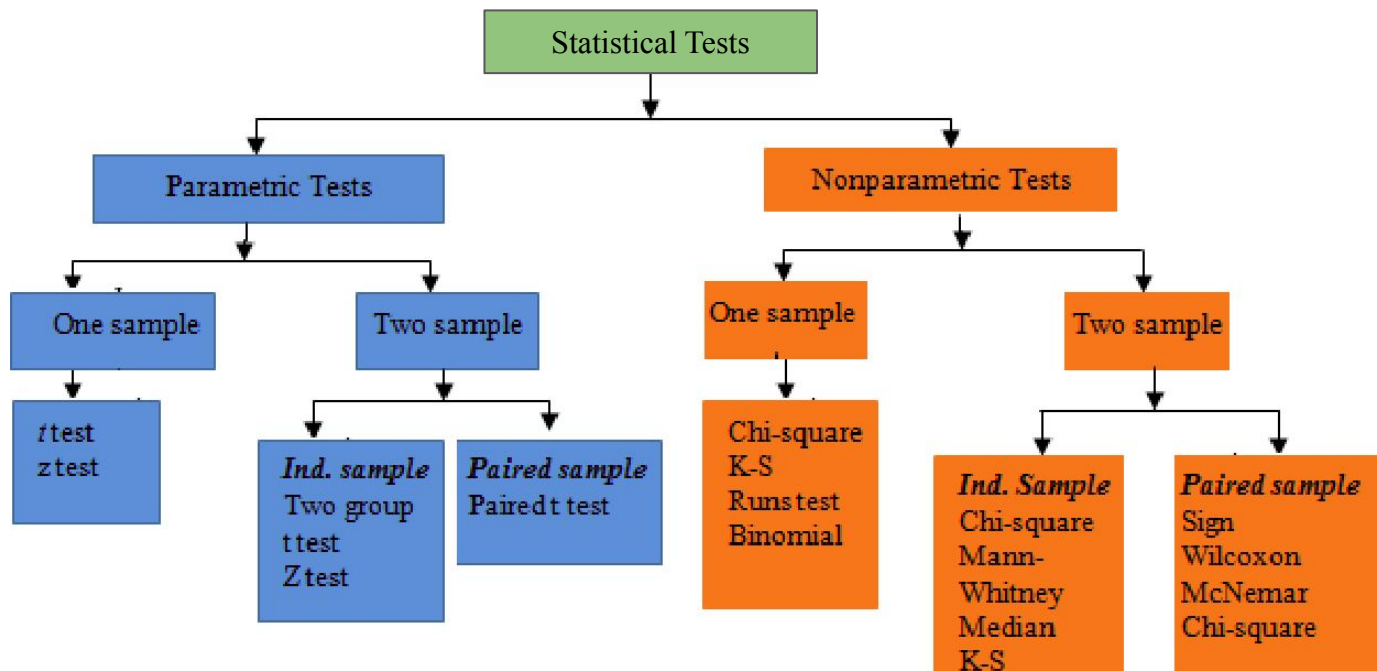


Tham khảo:

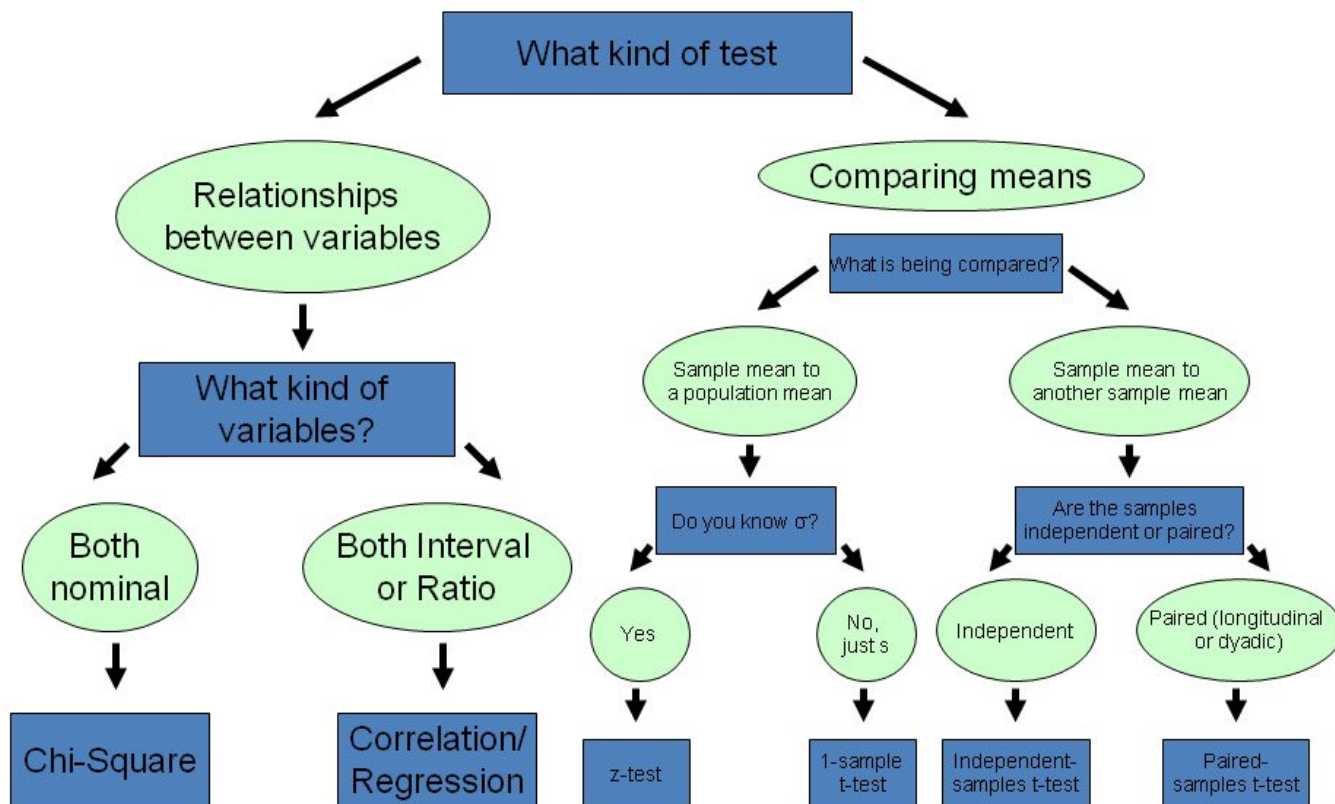
<https://www.analyticsvidhya.com/blog/2021/06/hypothesis-testing-parametric-and-non-parametric-tests-in-statistics/>

<https://cyberlab.edu.vn/>

Phân loại hai loại kiểm định chính (2):



Phân loại hai loại kiểm định chính (3):



Hai loại kiểm định chính:

- ❖ Với mỗi parametric test thường có một non-parametric test tương ứng
- ❖ Non-parametric test thường ổn định hơn (more robust) parametric test (Vì các kiểm định này đúng với nhiều tình huống hơn, tức cần ít điều kiện đúng hơn)

3.1 Parametric Test

(a) Z Test:

- ❖ Để xác định xem các giá trị means có khác nhau không
- ❖ Đã biết trước variance (hoặc std) của tổng thể
- ❖ Hoặc số lượng mẫu: $n > 30$
- ❖ Phân bố của quần thể là phân bố chuẩn

One sample Z-Test

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

\bar{x} = sample mean
 μ = population mean
 σ = population standard deviation
 n = sample size

Two sample Z-Test

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

3.1 Parametric Test

(a) Z Test:

- ❖ Để xác định xem các giá trị means có khác nhau không
- ❖ Đã biết trước variance (hoặc std) của tổng thể
- ❖ Hoặc số lượng mẫu: $n > 30$
- ❖ Phân bố của quần thể là phân bố chuẩn

3.1 Parametric Test

(a) Z Test: bảng phân phối Z-Test

- ❖ Tham khảo 1: <https://nghienctuugiaoduc.com.vn/bang-phan-phoi-chuan-z-z-distribution/>
- ❖ Tham khảo 2: <https://www.ztable.net/>

$$Z = \frac{x - \mu}{\sigma}$$

Z = standard score

x = observed value

μ = mean of the sample

σ = standard deviation of the sample

Standard Normal Probabilities

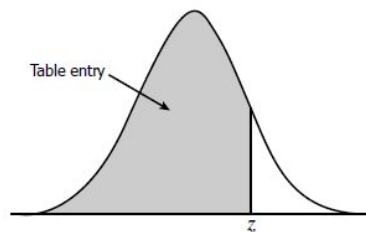


Table entry for z is the area under the standard normal curve to the left of z .

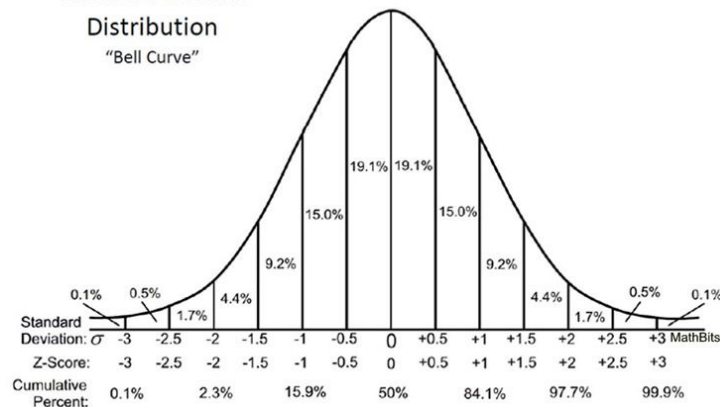
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177

3.1 Parametric Test

(a) Z Test: bảng phân phối Z-Test

$$Z = \frac{x - \mu}{\sigma}$$

Standard Normal
Distribution
"Bell Curve"



z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
+0.1	.53983	.54380	.54776	.55172	.55567	.55966	.56360	.56749	.57142	.57535
+0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
+0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
+0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
+0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
+0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
+0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
+0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
+0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
+1	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
+1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
+1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
+1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91308	.91466	.91621	.91774
+1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
+1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
+1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
+1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
+1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
+1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
+2	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
+2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
+2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
+2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
+2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
+2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
+2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
+2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
+2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
+2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
+3	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
+3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
+3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
+3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
+3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
+3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
+3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
+3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
+3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
+3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997
+4	.99997	.99997	.99997	.99997	.99997	.99997	.99998	.99998	.99998	.99998

Standard Normal Probabilities

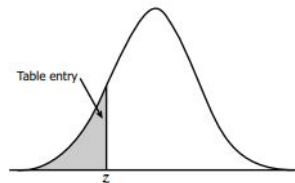


Table entry for z is the area under the standard normal curve to the left of z .

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Standard Normal Probabilities

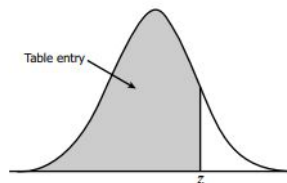


Table entry for z is the area under the standard normal curve to the left of z .

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

(b) T Test:

- ❖ Dựa trên phân bố student-T
- ❖ Không biết trước variance của tổng thể
- ❖ Số lượng mẫu: $n \leq 30$

One sample Z-Test

(So sánh sample mean với pop. mean)

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Two sample Z-Test

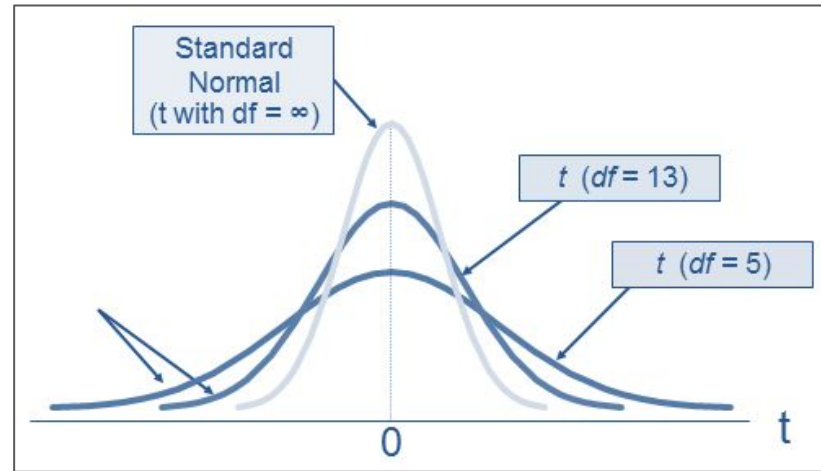
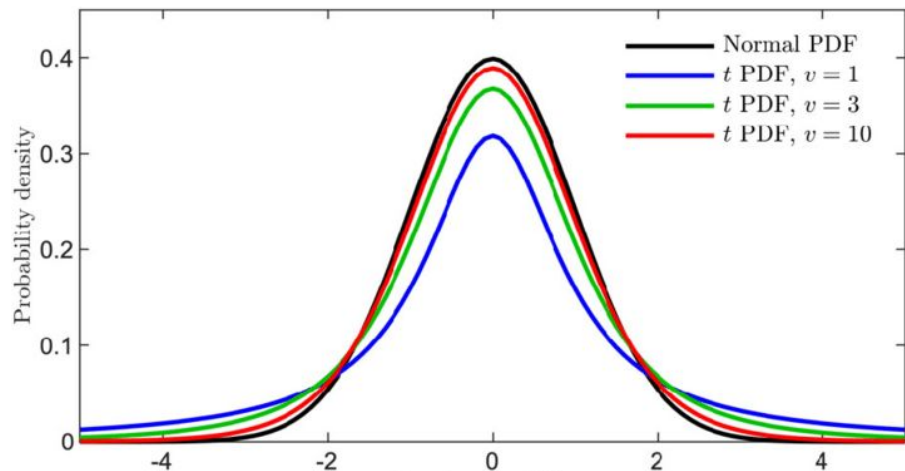
(So sánh means của hai samples khác nhau)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

3.1 Parametric Test

(b) T Test: bảng phân phối T-Test

❖ Tham khảo: <https://nghiencuugiaoduc.com.vn/bang-phan-phoi-t-t-distribution/>



3.1 Parametric Test

(b) T Test: bảng phân phối T-
Tham khảo:

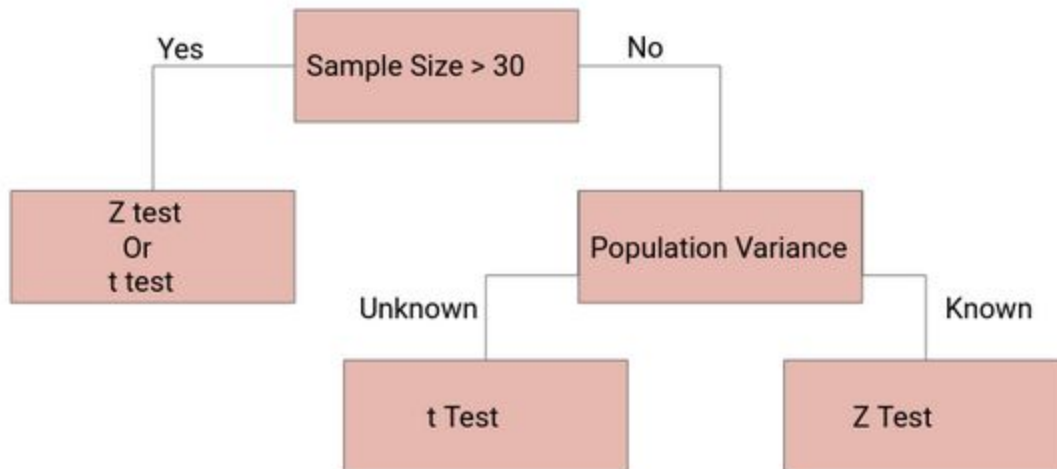
<https://www.tdistributiontable.com/>

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
Z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

3.1 Parametric Test

(a & b) T-Test vs. Z-Test



(c) F Test:

- ❖ Để xác định xem 02 quần thể có variance giống nhau hay không
- ❖ 02 quần thể này đều thuộc phân bố chuẩn (dựa trên Snedecor F-distribution)
- ❖ Còn gọi là So sánh mức độ giống nhau của sample variances
- ❖ F-statistic chỉ đơn giản là tỷ lệ giữa hai variance: $F = \text{var1}/\text{var2}$

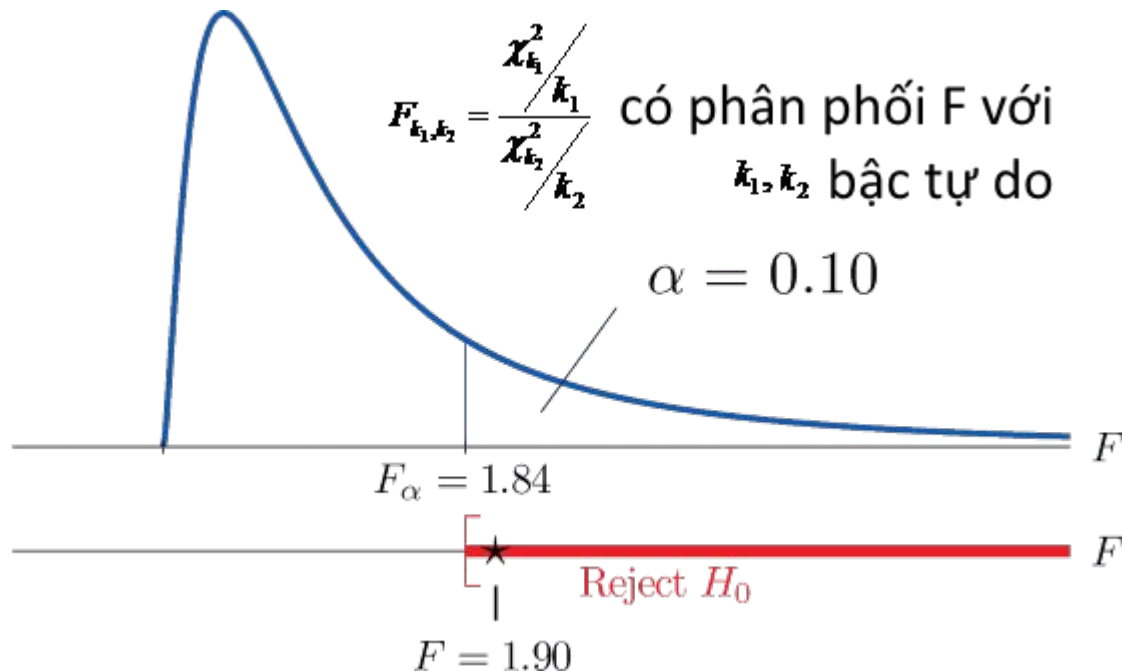
$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1}$$

$H_0: \sigma_1^2 = \sigma_2^2$ (the population variances are equal)

$H_1: \sigma_1^2 \neq \sigma_2^2$ (the population variances are *not* equal)

(c) F Test: bảng phân phối F-Test

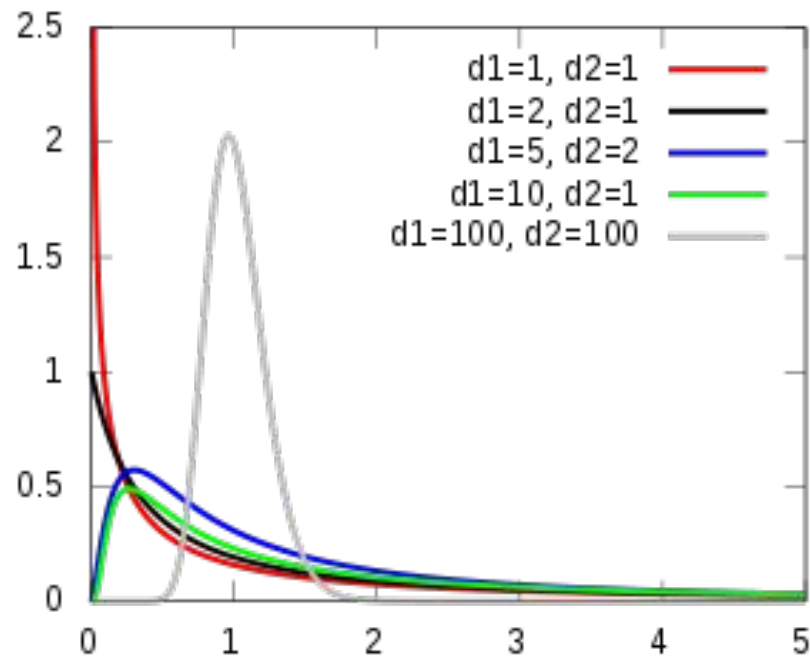
❖ Tham khảo: <https://nghiencuugiaoduc.com.vn/bang-phan-phoi-f-f-distribution/>



3.1 Parametric Test

(c) F Test: bảng phân phối F-Test

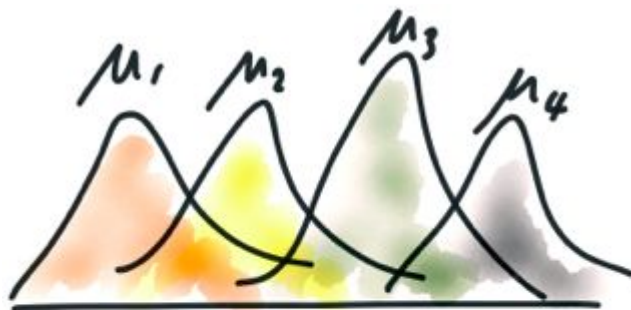
❖ Tham khảo: <https://en.wikipedia.org/wiki/F-distribution>



3.1 Parametric Test

(d) **ANOVA Test** (ANalysis Of VAriance, phân tích phương sai):

- ❖ Phân tích tổng quy mô sự biến thiên của biến phụ thuộc với một hay một nhóm các biến giải thích
- ❖ Dùng để so sánh means của nhiều groups độc lập thuộc cùng 1 cột dữ liệu



ANOVA

$$\mu_1 = \mu_2 = \mu_3 = \mu_4 ?$$

(d) ANOVA Test (ANalysis Of VAriance, phân tích phương sai):

❖ Phân tích tổng quy mô sự biến thiên

- Biến giải thích (nguyên nhân): thường là định tính (chia thành các nhóm)
- Biến phụ thuộc (kết quả): thường là định lượng
- Kiểm tra xem các mẫu thu được có phải từ cùng một tổng thể không, có tương quan với nhau không?
- Ví dụ:
 - Khu vực sống của học sinh có ảnh hưởng điểm thi ĐH môn tiếng Anh?
 - Giới tính khác nhau thì giá trị tổng đơn hàng có khác nhau không?

(d) ANOVA Test (ANalysis Of VAriance, phân tích phương sai):

❖ Tham khảo:

- Video (GS. Nguyễn Văn Tuấn):

<https://www.youtube.com/watch?v=I19vEdaM1eY&t=2s>

- Bài viết:

<https://hozo.vn/y-nghia-cua-kiem-dinh-anova/>

- Tiếng Anh:

<https://www.analyticsvidhya.com/blog/2020/06/introduction-anova-statistics-data-science-covid-python/>

(d) ANOVA Test:

❖ 3 loại kiểm thử ANOVA:

- One Way Anova: So sánh sample means của các groups trên một biến độc lập
- Two Way Anova: So sánh means (ảnh hưởng) của các groups thuộc 2 biến phân loại độc lập lên một biến phụ thuộc
- Manova (Multivariate Analysis of Variance): Có nhiều biến phụ thuộc

3.1 Parametric Test

(d) ANOVA Test:

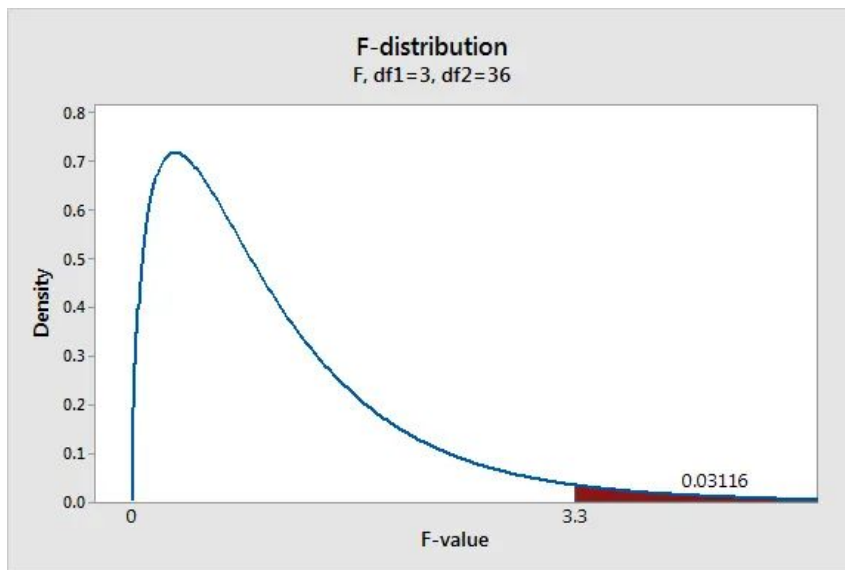
❖ ANOVA với F-Test

Tham khảo:

<https://statisticsbyjim.com/anova/f-tests-anova/>

$$F = \frac{\text{variance between treatments}}{\text{variance within treatments}}$$

$$F = \frac{MS_{\text{Treatments}}}{MS_{\text{Error}}} = \frac{SS_{\text{Treatments}} / (I - 1)}{SS_{\text{Error}} / (n_T - I)}$$



(d) ANOVA Test:

❖ ANOVA với F-Test

$$F = \frac{\text{Explained Variance}}{\text{Unexplained variance}} \text{ or } \frac{\text{Variance between groups}}{\text{Variance within groups}} = \frac{\text{Sum of squares between groups}}{\text{Sum of squares for error}}$$
$$= \frac{MST}{MSE} = \frac{\text{Mean squared error treatments}}{\text{Mean squared error}} = \frac{SS_B / Df_T}{SS_W / Df_E} = \frac{SS_B / (k-1)}{SS_W / (N-k)}$$

Where,

$$SS_B = \sum_i n_i (\bar{y}_i - \bar{y})^2$$

$$SS_W = \sum_{ij} (y_{ij} - \bar{y}_i)^2$$

Where,

y_i – sample mean in the i^{th} group
 n_i – number of observation in the i^{th} group
 \bar{y} – total mean of the data
 k – total number of the groups

y_{ij} – j^{th} observation in the out of k groups
 N – Overall sample size

3.2 Non-parametric Test

Chi-square Test:

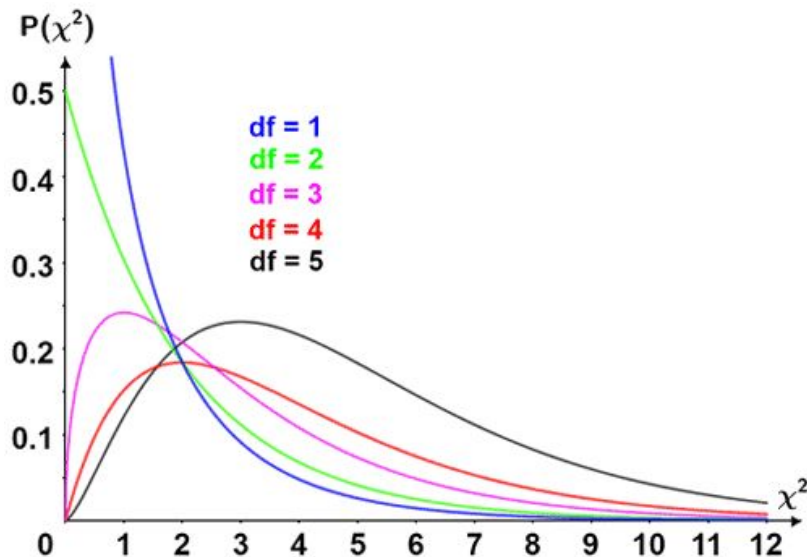
- ❖ Là một họ các kiểm thử phi tham số tuân theo phân bố Chi-square (χ^2)
- ❖ Thường dùng để:
 - Kiểm tra mức độ tốt của fitting
 - Kiểm tra mức độ độc lập (independence) của 02 biến

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O = the frequencies observed

E = the frequencies expected

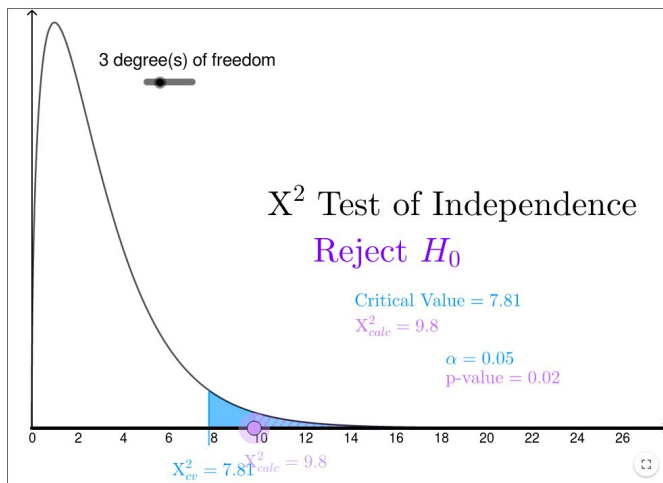
Σ = the 'sum of'



Chi-square Test:

- ❖ Thường gồm các kiểm thử:
 - Kiểm định chi bình phương Pearson
 - Kiểm định chi bình phương Yates
 - Kiểm định chi bình phương Mantel-Haenszel

3.2 Chi-square



Link thử nghiệm:

<https://www.geogebra.org/m/smhy8cxz>

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of χ^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

3.3 A/B Testing

Là trường hợp đặc biệt: Phần 4



Phần 4. A/B Testing

4.1. A/B Testing là gì?

4.2. Ví dụ

4.1 A/B Testing là gì?

- ❖ A/B Testing là một **phương pháp luận nghiên cứu** "trải nghiệm khách hàng"
 - Mục đích: customer experience
 - Còn gọi là **bucket testing** hoặc **split-run testing** (là cách để so sánh 2 versions của một "biến")
 - Có nền tảng là statistical hypothesis testing: **"02-sample hypothesis testing"**



Link: <https://www.invespcro.com/blog/ab-testing-statistics-made-simple/>

<https://cyberlab.edu.vn/>

4.1 A/B Testing là gì?

- ❖ Mục đích tiêu biểu:
 - Kiểm tra xem phản ứng (của khách hàng) với biến thể A so với biến thể B
 - Xác định xem biến thể/phiên bản nào hiệu quả hơn

- ❖ Ví dụ: một cty muốn thay đổi cách hiển thị thông tin sản phẩm trên website
 - Cách hiển thị mới có thu hút được nhiều khách hàng hơn?
 - Cách hiển thị mới có giúp tăng conversion rate (tỷ lệ khách hàng chọn mua)?
 - Các phiên bản/biến thể (variants):
 - **Phiên bản A** (control group): website cũ
 - **Phiên bản B** (treatment group): website mới

Lưu ý:

- Mỗi người dùng (trong thử nghiệm) chỉ nhìn thấy MỘT phiên bản
- An A/B test should have a defined outcome that is measurable such as number of sales made, click-rate conversion, or number of people signing up/registering

4.1 A/B Testing là gì?

Các tổ chức/tập đoàn lớn đã sử dụng large-scale A/B Testing: Airbnb, Amazon, Booking.com, Facebook, Google, LinkedIn, Lyft, Microsoft, Netflix, Twitter, Uber, Yandex, and Stanford University
⇒ The challenges can be grouped into four areas: Analysis, Engineering and Culture, Deviations from Traditional A/B tests, and Data quality.

Ví dụ

Google

- Ran their first A/B test in 2000: to determine what the optimum number of results to display on its search engine results page would be
- In 2011: Google ran over 7,000 different A/B tests

Microsoft

- Microsoft Bing (2000): test different ways of displaying advertising headlines
- Within hours, the alternative format produced a revenue increase of 12% with no impact on user-experience metrics

Today, companies like Microsoft and Google each conduct over 10,000 A/B tests annually (*designed experiment*).

THANK YOU!

