FIB

Departament d'Estadística
i Investigació Operativa

UNIVERSITAT POLITÈCNICA DE CATALUNYA

UPC

Inferència estadística

Bloc 4 – Probabilitat i Estadística

Novembre 2019

Índex

Inferència estadística. Guió

Guió de la part d'Estadística de PE:

- **B4**: Tècnica general de la inferència estadística
  - Estimar un paràmetre (***Intervals de Confiança***)
  - Refutar un paràmetre (***Proves d'Hipòtesis***)
- **B5**: Aplicació (I): Avaluació de millores
  - ***Disseny d'experiments:*** comparació de dues poblacions.
- **B6**: Aplicació (II): Predicció
  - ***Previsió*** d'una variable resposta en funció d'una variable explicativa.

Inferència estadística. Introducció

- **La ciència vol ser refutable:** El criteri de refutabilitat sosté que per ser col·locats en el rang de científics, els enunciats han de poder entrar en conflicte amb observacions possibles. [Ex: "Els marcians existeixen" : no hi ha mitjans per contradir-ho → no és científic]

- **La tècnica vol ser documentable:** S'ha de aportar evidència basada en dades [Ex: "El meu programa funciona": si no aportes proves/dades → no és tècnic]

- La metodologia estadística permet la inducció: **inferir les característiques de la població a partir de les observacions d'una mostra** [Ex: Per conèixer la velocitat mitjana real de connexió a tota la població amb un determinat proveïdor recullo una mostra de velocitats]

Inferència estadística. Introducció

- La Inferència Estadística defineix i **quantifica els riscos** d'aquest procés [Ex: No es pot conèixer la mitjana de la vel. de connexió a tota la població a no ser que es tingui dades de tota la població, però la estadística em permet quantificar l'error en l'estimació en una mostra concreta]
- Mètode científic i tècnic (estadístic):
  - per **deducció** → disseny de la recollida de dades (Població → Mostra)
  - per **inducció** → inferir (estimar) resultats (Mostra → Població)

Exemples:

- Vull dir "El meu programa funciona bé"
  - → 1) recollida de dades (*proves o evidència*)
  - → 2) anàlisi: estimar una mesura (ex: mitjana del rendiment)
- "El meu programa millora els resultats de ..."
  - → 1) recollida de dades (*proves o evidència*)
  - → 2) anàlisi: poder refutar la igualtat d'una mesura (ex: mitjana de rendiments)

Inferència estadísta. Mostra Aleatòria Simple (MAS)

Sigui la VA: $X: \Omega \rightarrow R$

$\omega_i \rightarrow X(\omega_i) = x_i$

Direm que M.A.S. de grandària **n** de la v.a. X

és una funció vectorial $M = (X_1, X_2, \ldots, X_n)$ tal que

$M: \Omega^n \rightarrow R^n$

$\omega = (\omega_1, \omega_2, \ldots \omega_n) \rightarrow M(\omega_i) = (X_1, X_2, \ldots, X_n)$

Direm que és una MAS si i només si es compleixen les dues condicions següents:

(1) **Tots els elements** de la població tenen la **mateixa probabilitat** de pertànyer a la mostra.

(2) **Qualsevol combinació** de *n* elements té la **mateixa probabilitat** de pertànyer a la mostra.

La informació aportada per les diferents unitats ha de ser **independent** entre sí:

- les X_i han de ser VA independents i idènticament distribuïdes (i.i.d.)

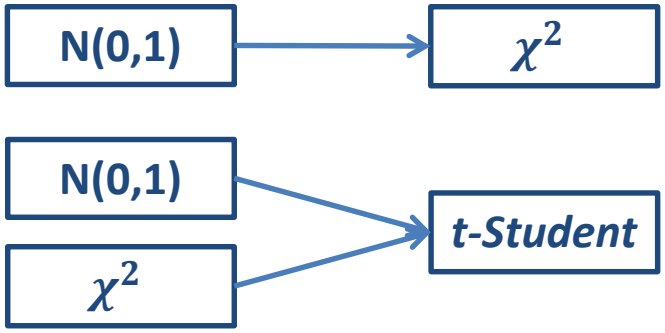
Inferència estadística. Conceptes bàsics

- **Paràmetre**: Indicador de la <u>població</u> que estem interessats en conèixer o estimar. [Ex: La μ (esperança) de les alçades dels estudiants de la FIB]
- **Estadístic**: qualsevol indicador que s'obtingui com a funció de les dades d'una <u>mostra</u>. [Ex: La suma de les alçades dels estudiants recollits en una mostra]
- **Estimador**: estadístic d'una <u>mostra</u> que s'utilitza per conèixer el valor d'un paràmetre de la <u>població</u>. [Ex: La mitjana de les alçades en una mostra d'alumnes de la FIB és una estimador de la μ (esperança) de les alçades dels estudiants de la FIB]

Atenció: La paraula ***mitjana*** pot voler dir ***paràmetre esperança*** quan parlem del centre de gravetat d'una distribució poblacional, o ***estadístic mitjana*** quan ens referim al promig d'una sèrie de valors obtinguts d'una mostra.

Models derivats de la Normal: χ^2 i *t-Student*

- Hi ha un parell de distribucions noves que serviran per abordar el contingut d'aquest tema: χ^2 i **t-Student**
- Aquestes distribucions provenen de fer operacions amb VA provinents d'altres distribucions, entre elles la Normal estàndard.



- A diferència de les distribucions vistes prèviament NO modelen fenòmens de la vida real, sinó el comportament dels estadístics entre les possibles mostres.

Distribució χ^2 (chi-quadrat)

- **Definició**: Siguin $X_i \sim N(0,1)$. Llavors:

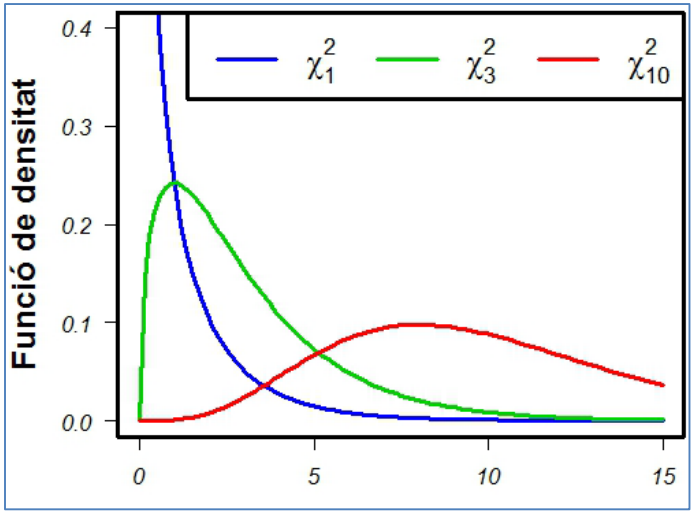
  $$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi_n^2$$

  [Concretament, per n = 1 → $X_1^2 \sim \chi_1^2$]

- **Notació**: $X \sim \chi_n^2$
- **Paràmetres**: *n* (graus de llibertat)
- **Funció de probabilitat i distribució**:

$$f(x) = \frac{x^{k/2-1} \cdot e^{-x/2}}{2^{k/2} \cdot \Gamma(k/2)} \quad per\ x > 0$$

$$F(x) = \frac{\gamma(k/2, x/2)}{\Gamma(k/2)} \quad per\ x > 0$$



Γ : funció Gamma
γ : funció Gamma incomplerta
n: graus de llibertat

R: dchisq, pchisq, qchisq

Script per veure que la suma de Normals estàndard al quadrat és una χ^2

```
M = 500                                          # Mostres de normals
n = 7                                            # Graus de llibertat
sample = array(rnorm(M*n),dim=c(M,n))            # n mostres de N(0,1)
sample2 = sample*sample                          # n mostres de (N(0,1))^2
sum = apply(sample2, 1, sum)                     # Suma de les mostres al^2
hist(sum, breaks="Scott", freq=FALSE)            # Distribució empírica sumant Normals
curve(dchisq(x, n),add=TRUE,col=2,lwd=2)# Distribució teòrica de la chi-quadrat
quantile(sum,c(0.25,0.50,0.75))                  # Q1, Mediana i Q3 de la suma de Normals
qchisq(c(0.25, 0.50, 0.75),n)                    # Q1, Mediana i Q3 de la chi-quadrat
```

Distribució t-Student

- **Definició**: Siguin dues VA independents, $Z \sim N(0,1)$ i $Y_n \sim \chi^2_n$. Llavors:

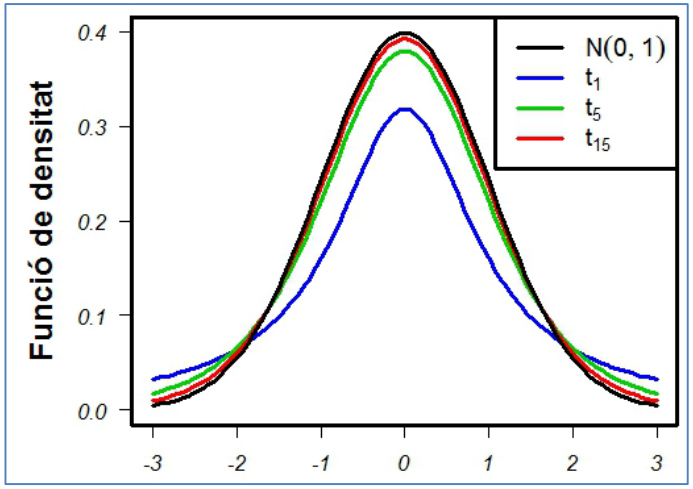
$$\frac{Z}{\sqrt{Y_n/n}} \sim t_n$$

[Quan n→∞ (n>30), llavors $t_n \rightarrow N(0,1)$]

- **Notació**: $X \sim t_n$
- **Paràmetres**: *n* (graus de llibertat)
- **Funció de probabilitat i distribució**:

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} \quad per\ x > 0$$

$$F(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n}\cdot B(1/2,n/2)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} \quad per\ x > 0$$



Γ : funció Gamma
B: funció Beta
n: graus de llibertat

R: dt, pt, qt

Script per veure que a partir de una Z i una Y_n s'obté una t

```
M = 500; n = 7                              # Número de mostres i graus de llibertat
samplez = rnorm(M, 0, 1)                    # Mostra de normals
samplechi2 = rchisq(M,n)                    # Mostra de chi-quadrats
samplechi2n = sqrt(samplechi2/n)            # Càlcul dels denominadors
t = samplez / samplechi2n                   # Càlcul de la t-student
hist(t, breaks="Scott", freq=FALSE)         # Distribució empírica
curve(dt(x, n),add=TRUE,col=2,lwd=2)        # Distribució teòrica d'una t-Student
quantile(t, c(0.25, 0.50, 0.75))            # Q1, Mediana i Q3 de Z/sqrt(Yn/n)
qt(c(0.25, 0.50, 0.75),n)                   # Q1, Mediana i Q3 de la chi-quadrat
```

Estimació puntual

- Un estimador $\widehat{\theta}$ del paràmetre desconegut θ, a partir de la mostra $M(\omega_i)$ $(X_1,X_2,\ldots,X_n)$ és una funció de les VA:

$$\hat{\theta} = f(X_1, X_2, \ldots, X_n)$$

- **Estimació puntual:** valor que l'estimador $\hat{\theta}$ pren en una mostra concreta. Ex: $\bar{x} = \frac{\sum x_i}{n}$ és la mitjana mostral i és una estimació puntual de μ

Nota: Distingiu entre el valor $\bar{x}$ i la variable aleatòria mitjana mostral $\bar{X}$ que no s'extreu de cap mostra

- **Error tipus o error estàndard:** variabilitat de l'estimador. Ex: en el cas anterior de la MITJANA, l'**error tipus (o estàndard) de la mitjana** (o ***mean standard error*** o **se**) és:

$$se = \sqrt{V(\bar{X}_n)} = \sqrt{E[(\bar{X}_n - \mu)^2]} = \frac{\sigma}{\sqrt{n}}$$

Nota: Generalment, la **σ** serà desconeguda i l'error tipus l'haurem d'aproximar emprant l'estimador pertinent ($\hat{\sigma}$) amb les dades de la mostra: $\widehat{se} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \sqrt{\sum \frac{(x_i-\bar{x})^2}{n-1}} \cdot \frac{1}{\sqrt{n}}$

Estimació puntual. Exemples

Nota: per als paràmetres s'utilitzen lletres de l'alfabet grec

Paràmetre (θ) (**POBLACIÓ**)	Estimador ($\hat{\theta}$) (**MOSTRA**)
μ (esperança, mitjana poblacional)	$\bar{x}$ (mitjana mostral)
σ^2 (variància poblacional) σ (desviació tipus poblacional)	s^2 (variància mostral) s (desviació tipus mostral)
π (probabilitat)	p (proporció)

Exemple:

En 9 dies consecutius s'ha observat el nombre de terminals en una Universitat connectats a internet: 587, 470, 676, 451, 436, 672, 584, 697 i 408 [$\Sigma x = 4981$; $\Sigma x^2 = 2860855$]

```
nterm <- c(587,470,676,451,436,672,584,697,408)
```

Una estimació puntual del nombre esperat (μ) de terminals diaris connectats és:

`mean(nterm)` → $\bar{x} = 553.44$ o $\bar{x} = (\sum x_i)/n = 553.44$

Una estimació puntual de la desviació tipus (σ) del nombre de terminals connectats és:

`sd(nterm)` → $s = 114.0988$ o $s = \sqrt{(\sum(x_i - \bar{x})^2)/(n-1)} = 114.0988$

L'estimació de l'error tipus o variabilitat de la mitjana és:

`sd(nterm)/sqrt(length(nterm))` → $se = 38.03 = \sqrt{(\sum(x_i - \bar{x})^2)/(n-1)} \cdot 1/\sqrt{n} = 38.03$

Estimació puntual. Propietats dels estimadors

- Inevitablement, les estimacions puntuals **fallen** o, millor dit, com depenen de la mostra que "ens ha tocat", **fluctuen** (encara que usualment tan sols observem un valor)
- Les 2 obsessions de l'Estadística són:
  - **quantificar** els errors d'estimació
  - **minimitzar** aquests errors
- L'error tipus o típic informa de l'**error esperat** a l'equiparar el valor de l'estimador obtingut en l'estudi amb el valor del paràmetre poblacional.
- Com l'estimador és "qualsevol" estadístic que s'utilitzi amb fins inferencials, hem de definir les propietats que permeten definir els "millors".

Nota: *l'error exacte en una mostra concreta roman desconegut, podent ser inferior o superior que l'error típic o esperat.*

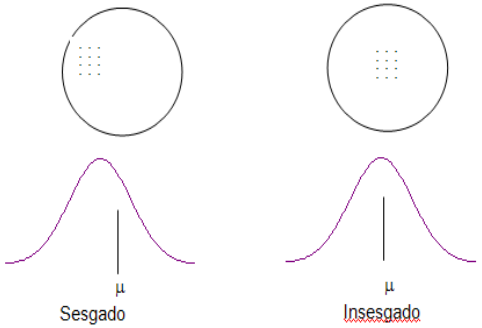
Estimació puntual. Propietats dels estimadors

Propietats desitjables

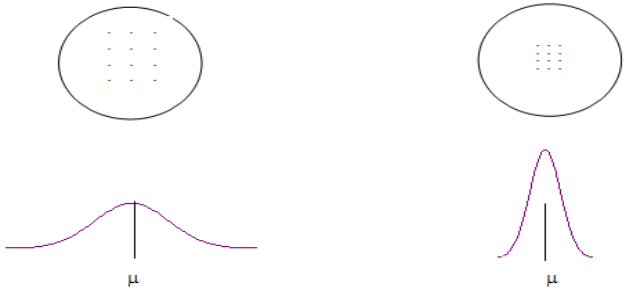
- **No tenir biaix** (= *sesgo*, *bias*) → $Biaix = 0$
  - Biaix és la diferència entre el centre de la distribució del estimador $[E(\hat{\theta})]$ i el valor del paràmetre a estimar $[\theta]$

$$Biaix = E(\hat{\theta}) - \theta$$

  - Un estimador $\hat{\theta}$ del paràmetre θ és NO esbiaixat si Biaix = 0

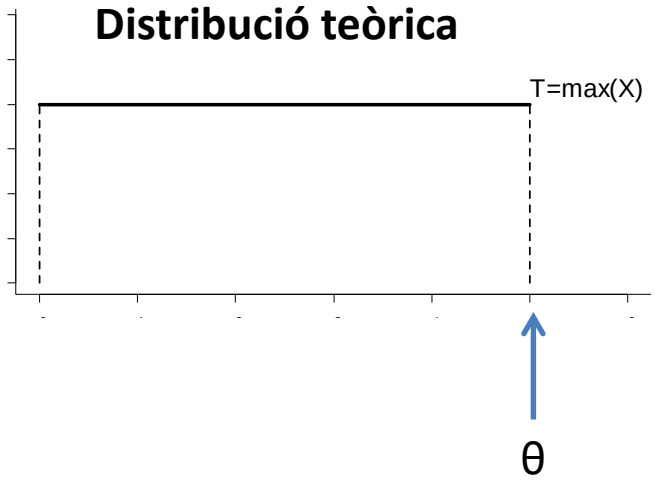


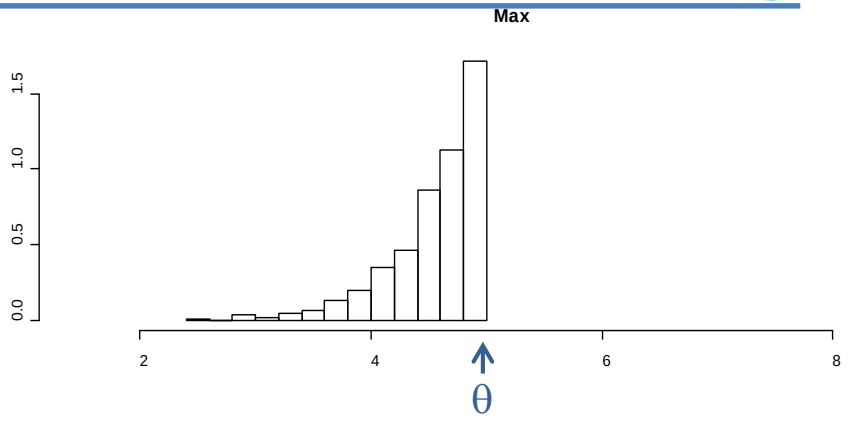
- **Ser Eficient** → $V(\hat{\theta}) \downarrow$
  - Entre dos estimadors NO esbiaixats, es diu que és més eficient el que té una variància menor.

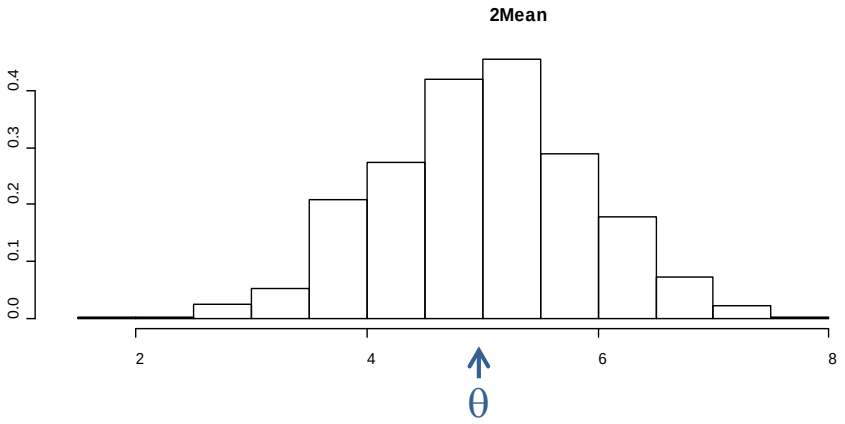


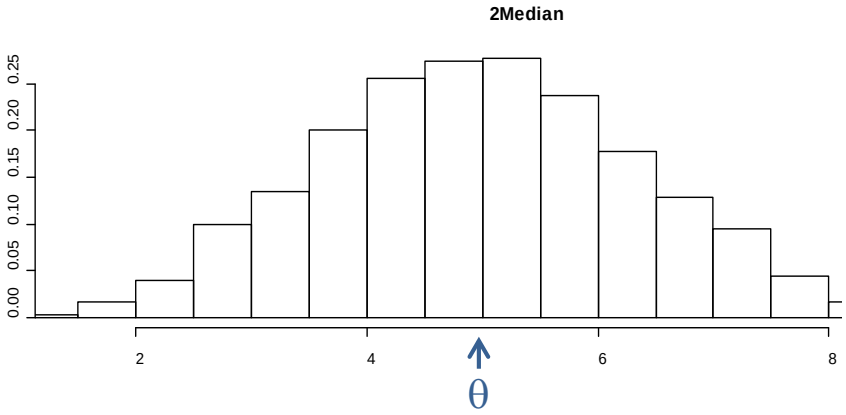
Estimació puntual. Exemple

Estimar el límit superior **θ** d'una llei uniforme, amb mínim zero: U[0 , **θ**]









Tres estimadors possibles:

- Màxim de la mostra (esbiaixat)
- Doble de la mitjana mostral (no esbiaixat)
- Doble de la mediana mostral (no esbiaixat)

Teorema del Límit Central (repàs)

- El T.L.C. estableix que, si s'agafen mostres de grandària *n* d'una població de mitjana μ i desviació típica σ, a mesura que creix *n*, la distribució de la mitjana mostral $\bar{X}$ s'aproxima a la d'una normal de mitjana μ i desviació típica $\sigma/\sqrt{n}$:

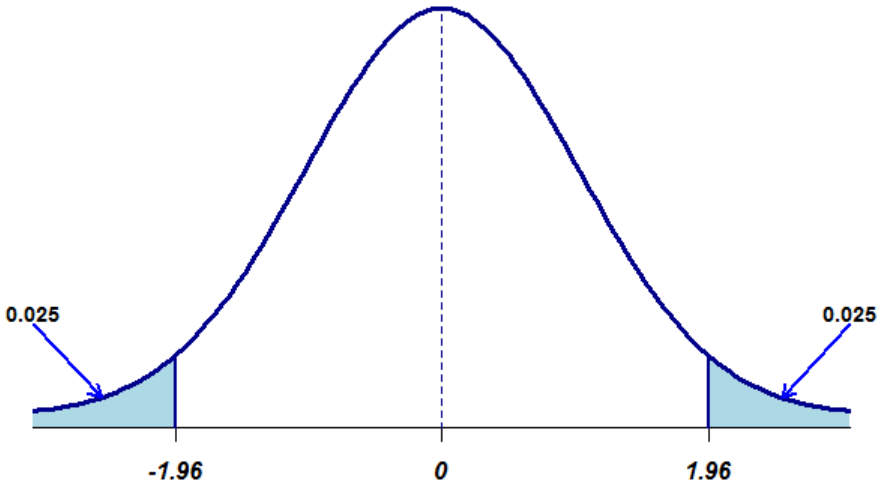
$$\bar{X}_n \xrightarrow[n\ gran]{} N(\mu, \frac{\sigma}{\sqrt{n}})$$

- ***¿Quina grandària ha de tenir n per a que es compleixi el TLC?***
  - Si X ~ N → $\bar{X}_n$~ N **∀n** (Les combinacions lineals de Normals i.i.d. són Normals)
  - Si X és quantitativa → $\bar{X}_n$ ~ N si **n > 30** (Com més s'assembli X a la Normal abans convergirà a una Normal)
  - Si X té una distribució **discreta** i/o **asimètrica**, la convergència requereix una grandària mostral (*n*) més gran

Quantils en la Z~N(0,1) (repàs)

Definició: El quantil α és el valor fins al qual s'acumula una probabilitat α

Notació: Aquell valor a tal que $F_Z(a) = P(Z \leq a) = \alpha$ l'indicarem per $\mathbf{z_\alpha}$



Exemples:

$\mathbf{z_{0.025} = -1.960}$ $\mathbf{z_{0.975} = 1.960}$ $\mathbf{z_{0.95} = 1.645}$

La Normal, al ser simètrica, compleix que:

$$z_\alpha = -z_{1-\alpha}$$

$$z_{\alpha/2} = -z_{1-\alpha/2}$$

Interval per a la mitjana mostral. Exemple

- Situació:

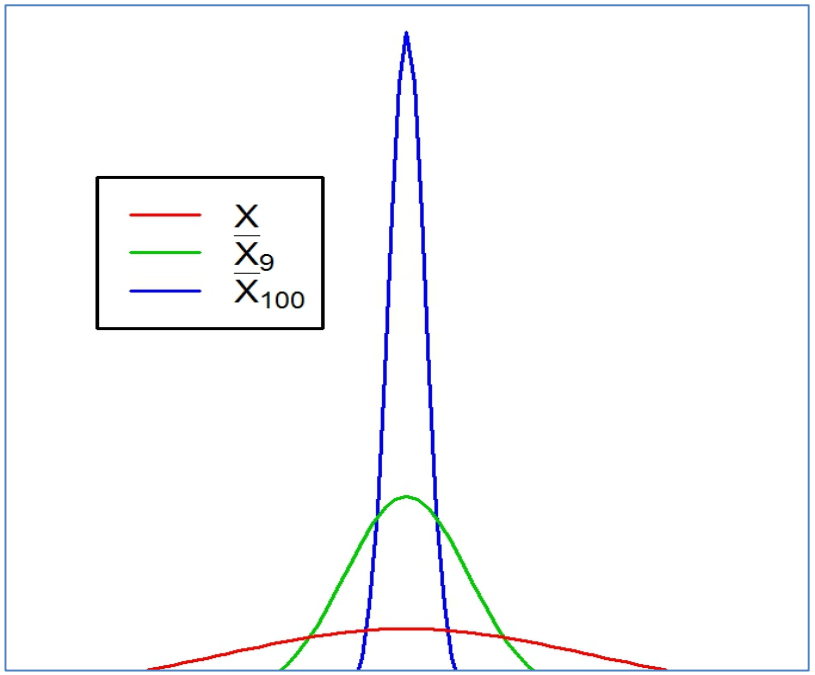
  X: "*temps d'execució d'un algoritme*" ~ N(μ = 100ms, σ = 10ms).

- Plantegem les distribucions de les VA: $X, \bar{X}_9 \; i \; \bar{X}_{100}$
- Calcularem, per *n*=1, *n*=9 i per *n*=100, els intervals amb probabilitats *grans* (95%) d'assegurar que les possibles $\bar{x}$ hi pertanyeran [deixem fora només una petita proporció (α: 0.05 o 5%)]
- Distribucions:

$$X = \bar{X}_1 \sim N(\mu = 100\ ms, \sigma = 10\ ms)$$

$$\bar{X}_9 \sim N\left(\mu = 100\ ms, \sigma = \frac{10}{\sqrt{9}} = 3.33\ ms\right)$$

$$\bar{X}_{100} \sim N\left(\mu = 100\ ms, \sigma = \frac{10}{\sqrt{100}} = 1\ ms\right)$$



Interval per a la mitjana mostral. Exemple

Els límits v, w dels intervals els podem calcular utilitzant les taules de la N(0,1):

$$z_{0.975} = 1.96 \qquad z_{0.025} = -1.96$$

- Rang que conté el 95% de les infinites execucions de l'algoritme (*X*)

$$v,w = \mu \pm z_{0.975} \cdot \sigma = 100 \pm 1.96 \cdot 10 = 100 \pm 19.60 = [80.40, 119.60]$$

- Rang que conté el 95% de les mitjanes de les infinites mostres de *n* = 9 execucions ($\bar{X}_9$)

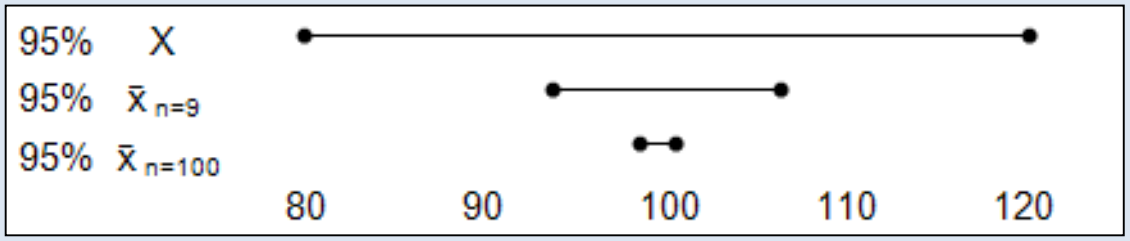
$$v,w = \mu \pm z_{0.975}\, \sigma/\sqrt{n} = 100 \pm 1.96 \cdot 10/3 = 100 \pm 6.53 = [93.47, 106.53]$$

- Rang que conté el 95% de les mitjanes de les infinites mostres de *n* = 100 execucions ($\bar{X}_{100}$)

$$v,w = \mu \pm z_{0.975}\, \sigma/\sqrt{n} = 100 \pm 1.96 \cdot 10/10 = 100 \pm 1.96 = [98.04, 101.96]$$

Representació esquemàtica:

Com l'amplada de l'interval depèn inversament de l'arrel de n, passar de n=1 a n=9, fa l'interval 3 vegades més estret



Estimació per interval de μ

- Hem vist que sabem calcular un "interval" que contingui $\bar{x}$ a partir de μ. Però el problema real és **calcular interval per μ, coneixent $\bar{x}$** (és a dir, passar d'un interval per a la mitjana mostral $\bar{x}$ a un per a la mitjana poblacional μ)
- A partir d'una probabilitat 1- α entre dos valors a i b (simètrics): *(amb σ coneguda)*

$$P(a \leq \bar{X}_n \leq b) = 1-\alpha \rightarrow P\left(\frac{a-\mu}{\sigma/\sqrt{n}} \leq \frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}} \leq \frac{b-\mu}{\sigma/\sqrt{n}}\right) = 1-\alpha \rightarrow P\left(z_{\frac{\alpha}{2}} \leq \frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}} \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

- Obtenim l'interval de la v. a. $\bar{X}_n$ amb **probabilitat** 1-α

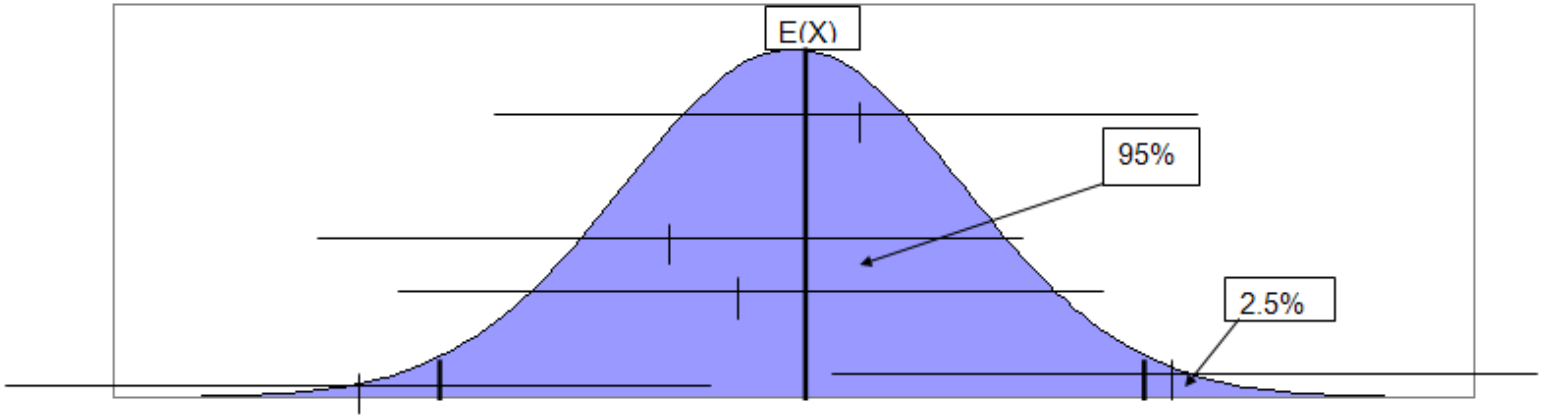
$$P\left(\mu + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \bar{X}_n \leq \mu + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

- I reordenant obtenim **l'interval de confiança 1-α del paràmetre μ**

$$P\left(\bar{X}_n + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Estimació per interval de μ

- $P\left(\bar{X}_n + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$ significa que amb una confiança de 1 - α podem assegurar que E(X) = μ estarà en el rang calculat.
- Si 1-α és 95% (α =5%): **el 95% dels intervals (IC) contindran μ**



- *Aquest procediment encerta el 100·(1-α)% de les vegades!*
- Denotem **IC(μ, 1-α)** a l'**INTERVAL DE CONFIANÇA** 1- α de μ, i l'expressem:

$$IC(\mu, 1-\alpha) = \bar{x} \mp z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

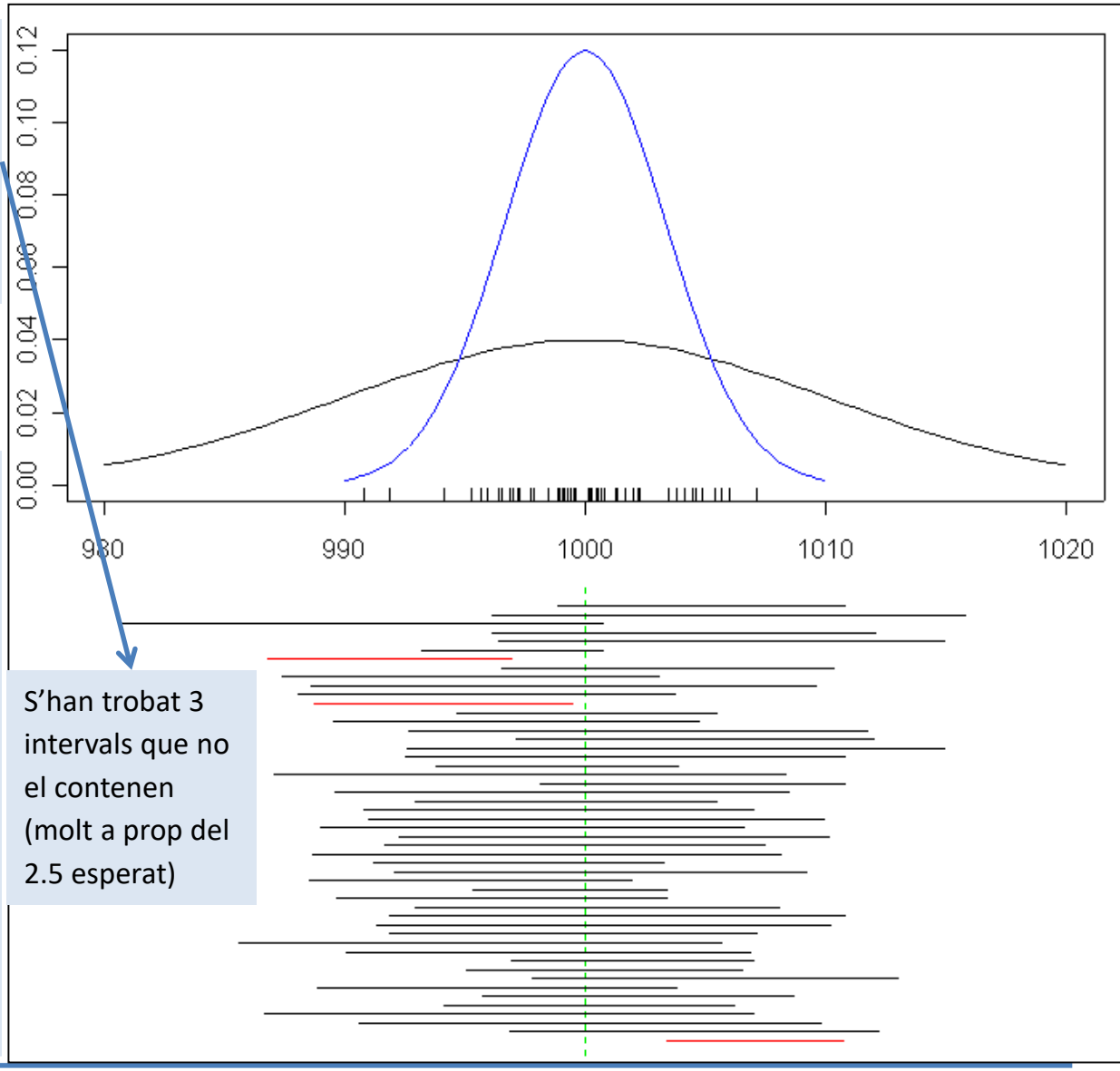
Com que la Normal és simètrica, llavors $z_{\alpha/2} = -z_{1-\alpha/2}$

Atenció: nosaltres només observarem una mostra, i no sabrem si l'IC trobat conté o no μ, però sí sabem que aquest procediment a la llarga dóna un 100·(1-α)% d'encerts.

Estimació per interval de μ. Simulació

S'han simulat 50 mostres amb n=9 provinents d'una N(μ=1000, σ). Calculant el IC95%, esperem que aproximadament el 5% (2.5) d'ells no continguin el valor real de μ

Nota tècnica: Amb un IC determinat (p.ex., [985, 1004]), s'ha de dir "**tenim un alt grau de confiança (p.ex., 95%) de que el paràmetre es trobi entre aquest dos valors**", però no és correcte parlar d'una probabilitat de 0.95 que el paràmetre estigui entre els dos valors trobats, perquè el paràmetre no es considera un element aleatori. Serà desconegut, però no és incert!



S'han trobat 3 intervals que no el contenen (molt a prop del 2.5 esperat)

Interval de confiança per μ (σ coneguda). Exemple

- Així doncs, l'interval de confiança 1−α de μ (amb σ coneguda) es calcula com:

$$IC(\mu, 1-\alpha) = \bar{x} \mp z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

- Recordeu que ens basem en el TCL i perquè es complís calia que la variable X inicial fos Normal o que n fos més gran de 30. Per tant, els requisits per realitzar aquest càlcul són: **n > 30 o X~N**

EXEMPLE:

Una embotelladora d'ampolles de litre té una dispersió de σ = 10cc. En una mostra a l'atzar de n = 100 ampolles d'aquesta màquina, la mitjana observada ha sigut $\bar{x}$ = 995cc. Calculeu un interval de confiança del 95% de μ.

$$IC(\mu, 0.95) = \bar{x} \mp z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 995 \mp 1.96 \cdot \frac{10}{\sqrt{100}} = 995 \mp 1.96 = [993.04\,, 996.96]$$

Resultat: amb una confiança del 95%, μ es troba entre 993.04 i 996.96

Interval de confiança per μ (σ coneguda). Exercici

1. La glicèmia en mmol/L té una desviació típica de σ = 1 en una mostra de n= 9 pacients, la mitjana $\bar{x}$ val 5. Calculeu el IC(μ,0.95).

   Amb una "força" del 95%, creiem que l'autèntic valor poblacional està entre aquests límits

2. Sense canviar la confiança, com podríem reduir l'interval a la meitat?

3. Calculeu l' IC amb una confiança del 99%

ATENCIÓ: quan *n* augmenta la precisió dels IC augmenta (interval més estret). Si augmenta la confiança (disminuint el risc *α* d'error), la precisió dels IC disminueix (interval més ample)

ATENCIÓ: En aquest cas, per estimar μ necessitem conèixer σ → *situació molt particular i infreqüent*

Interval de confiança. Mecànica

Passos	Esquema de Resolució
1	Definir l'estadístic a ser utilitzat
	Especificar la seva distribució
	Indicar les premisses necessàries per dir que segueix la distribució
	Delimitar el nivell de confiança (usualment $1-\alpha=95\%$)
2	Calcular l'interval
3	Interpretar el resultat

Interval de confiança per μ (σ desconeguda)

- Si **desconeixem σ**, la podem substituir per S, i llavors l'estadístic $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ passa a ser $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$ que és el quocient de 2 VAs i ja no es pot suposar que segueix una distribució N(0,1).
- Tal com diu el pas 2 de la mecànica de construcció de l'IC, cal conèixer la llei de probabilitats que segueix l'estadístic. En aquest cas, es tracta d'una nova distribució anomenada ***t-Student***.
- Per aquest estadístic, la distribució de probabilitat concreta és t_{n-1} (*n-1* graus de llibertat). Els percentils es poden trobar a taules específiques o amb R.
- Així doncs, l'interval de confiança 1-α de μ (amb σ desconeguda) és:

$$IC(\mu, 1-\alpha) = \bar{x} \mp t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$$

NOTA: la situació de no conèixer la σ de la població és més freqüent

Interval de confiança per μ (σ desconeguda)

- Podem utilitzar-la per conèixer el quocient informació/soroll utilitzant S en lloc de σ

$$\hat{t} = \frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}} \sim t_{n-1}$$

- Demostració:

$$\hat{t} = \frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}} = \frac{(\bar{X} - \mu)/\sqrt{\sigma^2/n}}{\sqrt{\frac{s^2/n}{\sigma^2/n}}} = \frac{Z}{\sqrt{\frac{s^2}{\sigma^2}}} = \frac{Z}{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)}{\sigma^2}}} = \frac{Z}{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2/\sigma^2}{n-1}}}$$

$$= \frac{Z}{\sqrt{\frac{Y_n}{n-1}}} \sim t_{n-1}$$

- Coneixem que el quocient informació/soroll segueix una 't' de Student .
- Això ens ajudarà a conèixer el IC95% per la μ amb σ desconeguda

Nota: Observeu que es requereix la Normalitat de les X's

Interval de confiança per μ (σ desconeguda). Ex.

En 9 dies consecutius s'ha observat el nombre de terminals en una Universitat connectats a internet: 587, 470, 676, 451, 436, 672, 584, 697 i 408.

R: `nterm <- c(587,470,676,451,436,672,584,697,408)`

Recordem que havíem calculat les estimacions puntuals: Una estimació del IC al 95% de confiança de la mitjana poblacional, assumint que la desviació poblacional (σ) val 100:

1-α	σ	IC(μ,1-α)	Resolució amb R

95%	Coneguda (σ=100)	[488.11; 618.78]	`n <- 9 ; sigma <- 100`
			`mean(nterm)- qnorm(0.975)*sigma/sqrt(n)`
			`mean(nterm)+ qnorm(0.975)*sigma/sqrt(n)`

99%	Coneguda (σ=100)	[467.58 ; 639.31]	`n <- 9 ; sigma <- 100`
			`mean(nterm)- qnorm(0.995)*sigma/sqrt(n)`
			`mean(nterm)+ qnorm(0.995)*sigma/sqrt(n)`

95%	Desconeguda	[465.74; 641.15]	`n <- 9`
			`mean(nterm)- qt(0.975,8)*sd(nterm)/sqrt(n)`
			`mean(nterm)+ qt(0.975,8)*sd(nterm)/sqrt(n)`

99%	Desconeguda	[425.83 ; 681.06]	`n <- 9`
			`mean(nterm)- qt(0.995,8)*sd(nterm)/sqrt(n)`
			`mean(nterm)+ qt(0.995,8)*sd(nterm)/sqrt(n)`

Interval de confiança per μ. Premisses

Per garantir el nivell de confiança de l'IC, s'ha de complir certes premisses

- Si sigma és coneguda, exigirem una de les condicions:
  - **X~N** → la combinació lineal de Normals és Normal ($\bar{X}$~N)
  - **Tenir una mostra gran (n≥30)** → Pel TCL , $\bar{X}$~N
- Si sigma no és coneguda, exigirem una de les condicions:
  - **X~N** → $(\bar{x}-\mu)/\sqrt{s^2/n} \sim t_{n-1}$
  - **Tenir una mostra gran (n≥100)** → Pel TCL , $\bar{X}$~N

La premissa que sempre s'ha de complir és que **l'origen de la mostra ha de ser aleatori** (v.a.i.i.d)

Dist. de referència	σ coneguda	σ desconeguda
X Normal	Normal	t de *Student*
X no Normal	Normal *si n ≥ 30*	Normal *si n ≥ 100*

Amb grans mostres la variació de "*s*" serà limitada (s estima molt bé σ), i podem considerar que

$$(\bar{x}-\mu)/\sqrt{s^2/n} \approx (\bar{x}-\mu)/\sqrt{\sigma^2/n} \sim \mathrm{N}(0,1)$$

[Aplicable al cas X~N, σ desconeguda i n gran]

Interval de confiança per π en una Binomial (n,π)

- Sigui X ~ B(n,π) → E(X) = π·*n*
  V(X) = π·(1-π)·*n*
- Sigui P = X/n → E(P) = E(X/*n*) = E(X)/*n* = π·*n* / *n* = π
  V(P) = V(X/*n*) = V(X)/n^2 = π·(1-π)·*n*/n^2 = π·(1-π)/*n*
- Per construir l'IC es pot recorre a la convergència de la Binomial a la Normal [amb la premissa de *n* gran i π no extrema → **π·*n* ≥ 5 i (1-π)·*n* ≥ 5**]:

$$P \to N\left(\mu_P = \pi, \sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}}\right)$$

- Així, l'interval de confiança s'assemblaria al de μ:

$$IC(\pi, 1-\alpha) = P \mp z_{1-\frac{\alpha}{2}}\sigma_P = P \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\pi(1-\pi)}{n}}$$

La **paradoxa** de que necessitem conèixer π per estimar el IC de π es pot solucionar de 2 maneres:

a) Substituint π per P→ $IC(\pi, 1-\alpha) = P \mp z_{1-\alpha/2} \cdot \sqrt{\big(P(1-P)\big)/n}$

b) Aplicant el màxim de π·(1-π) que correspon a π=0.5 → $IC(\pi, 1-\alpha) = P \mp z_{1-\alpha/2} \cdot \sqrt{\big(0.5(1-0.5)\big)/n}$

IC per π en una Binomial(n,π). Exemple

Llencem 100 vegades una moneda a l'aire i observem 56 cares (P = 56/100 = 0.56).

Les dues solucions per l'IC segons com estimem π:

$$IC(\pi, 0.95) = P \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{P\cdot(1-P)}{n}} = 0.56 \mp 1.96\sqrt{\frac{0.56\cdot 0.44}{100}} \approx 0.56 \mp 0.10 = [0.46,0.66]$$

$$IC(\pi, 0.95) = P \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\pi_m\cdot(1-\pi_m)}{n}} = 0.56 \mp 1.96\sqrt{\frac{0.50\cdot 0.50}{100}} \approx 0.56 \mp 0.10 = [0.46,0.66]$$

Donen el mateix IC fins al 2n decimal. El motiu és que la probabilitat estimada (0.56) és molt similar a la probabilitat de màxima indeterminació (0.50)

Interval de confiança per σ^2 . Distribució χ^2

- Gràcies a aquesta distribució, coneixem la distribució de S^2 (estimador de σ^2):

$$s^2 = \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n-1} \rightarrow (n-1)\cdot\frac{s^2}{\sigma^2} \sim \chi^2_{n-1}$$

- Demostració: si $X_i \rightarrow N$, llavors

Les x_i han de provenir d'una Normal

$$n\cdot\frac{\hat{\sigma}^2}{\sigma^2} = n\cdot\frac{\left(\sum_{i=1}^{n}(x_i-\mu)^2\right)/n}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 \sim \chi^2_n$$

- També, si $X_i \rightarrow N$

$$(n-1)\cdot\frac{\hat{\sigma}^2}{\sigma^2} = (n-1)\cdot\frac{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)/(n-1)}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2 \sim \chi^2_{n-1}$$

- *Per tant, **coneixem la distribució de s^2!!*** i podrem relacionar la distribució χ^2_{n-1} amb s^2 per definir IC de σ^2, tal com fem amb les distribucions Z i t_{n-1} amb $\bar{x}$ per definir IC de μ.

Interval de confiança per σ^2

- Hem vist que: $(n-1)\cdot\frac{S^2}{\sigma^2}\sim\chi^2_{n-1}$ sempre que $\mathbf{X_i \sim N}$ **(premissa)**
- Per tant:

$$P\left(\chi^2_{n-1,\frac{\alpha}{2}} \leq \frac{S^2\cdot(n-1)}{\sigma^2} \leq \chi^2_{n-1,1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$P\left(\frac{1}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \leq \frac{\sigma^2}{S^2\cdot(n-1)} \leq \frac{1}{\chi^2_{n-1,\frac{\alpha}{2}}}\right) = 1-\alpha$$

$$P\left(\frac{S^2\cdot(n-1)}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{S^2\cdot(n-1)}{\chi^2_{n-1,\frac{\alpha}{2}}}\right) = 1-\alpha$$

Atenció: és un IC per σ^2, no per σ!!

$$\boldsymbol{IC(\sigma^2, 1-\alpha) = \left[\frac{s^2(n-1)}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{s^2(n-1)}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]}$$

Nota: no és un interval simètric, ja que la distribució no ho és. Això implica calcular dos valors en la distribució en lloc de fer $\pm$.

IC per σ^2 en una Normal (μ,σ). Exemple

En les 25 execucions d'un mateix programa s'ha observat una variabilitat $s^2 = 8^2$

$$IC(\sigma^2, 0.95) = \left[\frac{s^2(n-1)}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{s^2(n-1)}{\chi^2_{n-1,\frac{\alpha}{2}}}\right] =$$

$$= \left[\frac{8^2(25-1)}{12.401}, \frac{8^2(25-1)}{39.364}\right] = [123.86, 39.02] \rightarrow \textbf{\textit{Oops! M'he equivocat}}$$

$$= \left[\frac{8^2(25-1)}{39.364}, \frac{8^2(25-1)}{12.401}\right] = [39.02, 123.86] \rightarrow \textbf{\textit{Ara sí!}}$$

Resultat:

$$IC(\sigma^2, 0.95) = [39.02, 123.86]$$

$$IC(\sigma, 0.95) = [6.25, 11.13]$$

Fent l'arrel quadrada, obtenim un interval per σ

Formulari per IC

TCL: $X_1,\ldots, X_n$ i.i.d. ($n\to\infty$), amb $E(X_i)=\mu$ i $V(X_i)=\sigma^2$,

llavors $\dfrac{\sum_{i=1}^n X_i}{n} = \bar{X}_n \approx N(\mu,\sigma^2/n)$ **i també** $\sum_{i=1}^n X_i \approx N(n\mu,\sigma^2 n)$ **)**

Estadístic mitjana mostral ($\bar{x}$): $\dfrac{(\bar{x}-\mu)}{\sqrt{\sigma^2/n}} \approx N(0,1)$ $\dfrac{(\bar{x}-\mu)}{\sqrt{s^2/n}} \approx t_{n-1}$ on $\bar{x} = \sum_{i=1}^n x_i / n$

Estadístic variància mostral (s^2): $s^2\dfrac{n-1}{\sigma^2} \approx \chi^2_{n-1}$ on $s^2 = \dfrac{\sum_{i=1}^n (x_i-\bar{x})^2}{n-1} = \dfrac{\sum_{i=1}^n x_i^2 - n(\bar{x})^2}{n-1} = \dfrac{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}}{n-1}$

Paràmetre	Estadístic	Premisses	Distribució	Interval de Confiança 1-α (Risc α)
μ	$\hat{z} = \dfrac{(\bar{x}-\mu)}{\sqrt{\sigma^2/n}}$	[X → N ò n ≥ 30] i σ coneguda	$\hat{z} \to N(0,1)$	$\mu \in (\bar{x} \pm z_{1-\alpha/2}\sqrt{\dfrac{\sigma^2}{n}})$
μ	$\hat{t} = \dfrac{(\bar{x}-\mu)}{\sqrt{s^2/n}}$	X → N	$\hat{t} \to t_{n-1}$	$\mu \in (\bar{x} \pm t_{n-1,1-\alpha/2}\sqrt{\dfrac{s^2}{n}})$
μ	$\hat{z} = \dfrac{(\bar{x}-\mu)}{\sqrt{s^2/n}}$	n ≥ 100	$\hat{z} \to N(0,1)$	$\mu \in (\bar{x} \pm z_{1-\alpha/2}\sqrt{\dfrac{s^2}{n}})$
σ (normal)	$\hat{X}^2 = \dfrac{s^2(n-1)}{\sigma^2}$	X → N	$\hat{X}^2 \to \chi^2_{n-1}$	$\sigma^2 \in \left(\dfrac{S^2(n-1)}{\chi^2_{n-1,1-\alpha/2}}, \dfrac{S^2(n-1)}{\chi^2_{n-1,\alpha/2}}\right)$
π (Binomial)	$\hat{z} = \dfrac{(p-\pi)}{\sqrt{\pi(1-\pi)/n}}$	(1-π)n ≥ 5 πn ≥ 5	$\hat{z} \to N(0,1)$	$\pi \in (P \pm z_{1-\alpha/2}\sqrt{\dfrac{\hat{\pi}(1-\hat{\pi})}{n}})$ $\hat{\pi} = P \quad o \quad \hat{\pi} = 0'5$
λ (Poisson)	$\hat{z} = \dfrac{(L-\lambda)}{\sqrt{\lambda}}$	λ ≥ 5	$\hat{z} \to N(0,1)$	$\lambda \in (L \pm z_{1-\alpha/2}\sqrt{L})$

El color gris indica un IC secundari

Proves d'hipòtesi. Exemple

Afirmo que encerto el 80% dels meus tirs lliures a basket, i un amic em posa a prova. Dels 20 tirs, solament faig 8. *"Fuà! Algú que encerta un 80% gairebé mai faria 8 de 20; així que no em crec la teva afirmació".*

El raonament de l'amic es basa en demanar-se què passaria si l'afirmació fos certa i es repetís la mostra de 20 tirs moltes vegades. Segurament poques vegades encertaria un nombre tan baix com 8. Un resultat com aquest és tan poc freqüent que aporta certa *evidència* per rebutjar la meva afirmació.

De fet, **aquesta *evidència* es pot quantificar**:

$$\textbf{Sigui M = nº encerts: } M \sim B(20, 0.8) \rightarrow P(M \leq 8) = 0.000102$$

Es a dir, si jo fos tan bo, solament 1 de cada 10000 vegades obtindria una puntuació tan nefasta (o pitjor).

Dues opcions: o he tingut mala sort o l'afirmació era falsa!

Source: The Basic Practice of Statistics. David S. Moore. 4th Ed.

Proves d'hipòtesi. Raonament

- Al igual que amb els IC, hem de pensar com si l'experiència es pogués repetir un gran nombre de vegades.
- Però ara no volem l'IC que contengui el paràmetre, ara partim d'una afirmació (una **hipòtesi** de partida, o **nul·la**), i volem estudiar si les dades proporcionen proves en contra seu.
- Una repetició intensa (una mostra infinita = la població) seria definitiva.
- Però, amb una mostra finita, quina informació aporten les dades?

Nota tècnica:

- Formalment, es distingeix entre les proves de Fisher (per aportar coneixement o evidencia o inferència) i els contrastos de Neyman-Pearson (per minimitzar els errors al prendre decisions). Els primers son rellevants per la Ciència (p.e., la Física) i els segons per la Tècnica (p.e. la Arquitectura). Però en aquest curs no distingirem i ho englobarem tot sota Proves de Hipòtesi (PH).
- Read more in: http://onlinestatbook.com/chapter9/significance.html

Proves d'hipòtesi. Raonament

La hipòtesi nul·la (H_0) es planteja formalment amb un paràmetre (o varis). El paràmetre en qüestió pren un valor que representem:

$$H_0: \pi = 0.80$$

π representa la probabilitat poblacional d'encertar un tir lliure, i volem comprovar si aquest valor és coherent amb les observacions.

Al igual que amb els IC, la mostra es *concentra* en un estadístic, que segueix una distribució de probabilitat coneguda si s'assumeix certa la H_0.

Addicionalment a H_0 afegim la hipòtesi **alternativa** H_1, que pot ser totalment complementària a la nul·la (enfoc bilateral), o parcialment (unilateral):

$$H_1: \pi \neq 0.80 \qquad\qquad H_1: \pi < 0.80$$

H_1 determina el(s) sentit(s) més oposat(s) a H_0: per exemple, el nombre de encerts pot ser l'estadístic, i si H_1 fos "≠" serien *sospitosos* tant els nombres d'encerts que van cap a 0 com els que van cap a 20. Si H_1 fos "<" serien *sospitosos* només els que van cap a 0 (que és el que hem pres, donat que el meu amic no confia molt en les meves habilitats).

Proves d'hipòtesi. P-valor

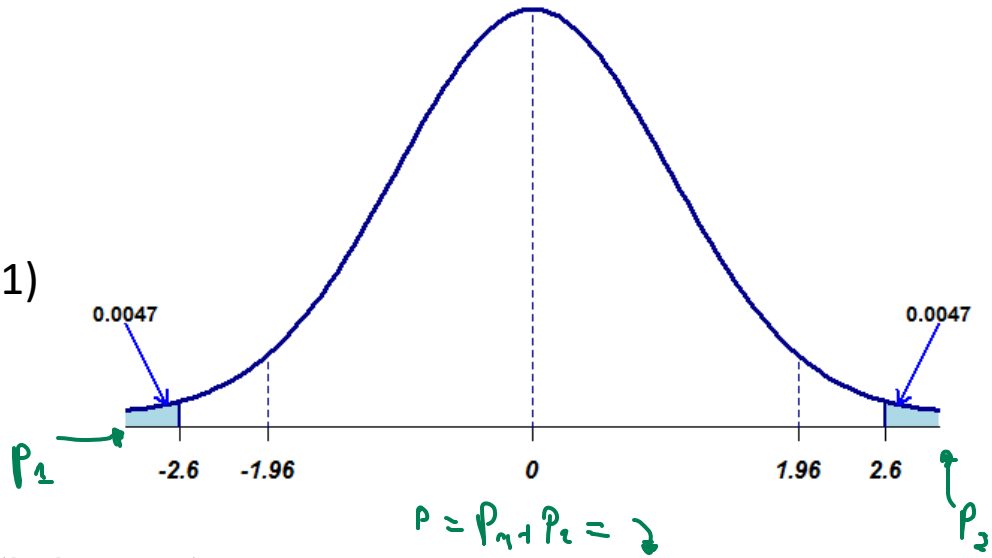
P, o **p-valor**, és la probabilitat de, sota H_0, obtenir resultats igual o més *extrems* que el observat.

Exemple:

Estadístic: Z = 2.6

Distribució de l'estadístic (sota H0): Z ~ N(0,1)

Contrast: bilateral



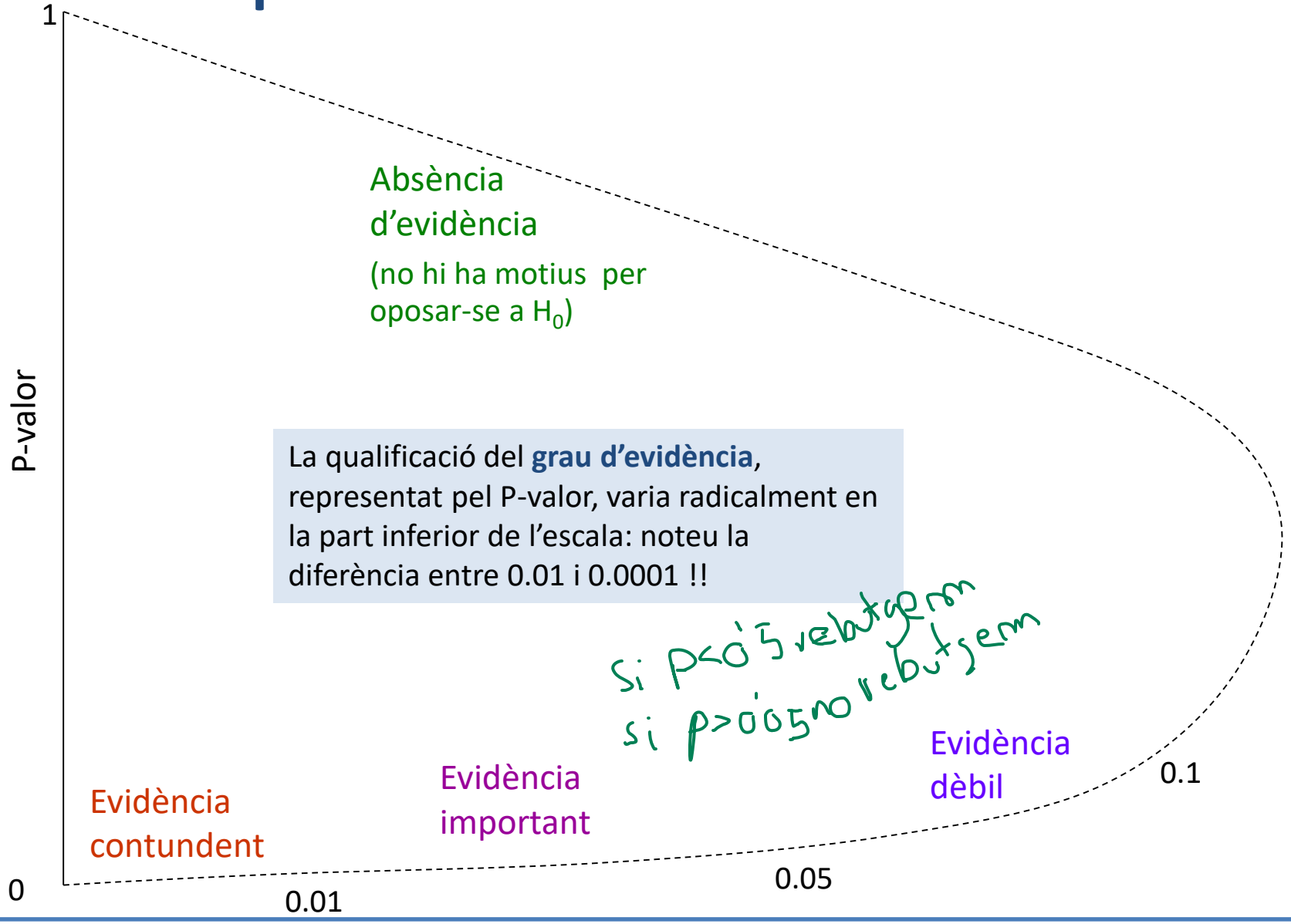
$$\left.\begin{array}{l} P(Z < -2.6) = 0.0047 \\ P(Z > +2.6) = 0.0047 \end{array}\right\} \rightarrow p - valor = P(|Z| > 2.6) = 2 \cdot 0.0047 = 0.0094$$

[**Taules**: 2·(1-0.9953) ; **R**: pnorm(-2.6) + (1-pnorm(2.6)))

P = p-valor = P(VA Z "*més lluny*" de H_0 que el valor observat z)

RECORDEU: "sota H_0" = suposem ('temptativament') que és cert que els paràmetres poblacionals valen el que diu H_0

Proves d'hipòtesi. P-valor



Malentessos amb el p-valor

- El *P*-valor diu amb quina freqüència poden passar events com el de la mostra (o més extrems) quan la hipòtesi H_0 és correcta:
  - Si el *P-valor* és petit $\rightarrow$ tenim evidència en contra de H_0
  - Si el *P*-valor no és petit, **NO** demostra la "veritat" de H_0
- *Incorrecte: P* = 0.000 $\rightarrow$ Correcte: *P*<0.001.
- *P-valor* **NO** és cap de les següents probabilitats:
  - la probabilitat d'haver-se equivocat
  - la probabilitat que la hipòtesi nul·la sigui certa
  - la probabilitat d'haver rebutjat erròniament la hipòtesi nul·la
- 1–*P*-valor **NO** és la probabilitat que la hipòtesi alternativa sigui certa
- Trobareu més a "Frequent misunderstandings" a Wikipedia

Proves d'hipòtesi. Resolució

1. Escollir una **variable** segons els objectius de l'estudi
2. Escollir un disseny i un **estadístic**
3. Definir una **hipòtesi** H_0 que es vol posar a prova, enfront una hipòtesi alternativa H_1
4. Especificar la **distribució** de l'estadístic si H_0 fos certa (i les premisses adients)
5. Amb les dades, calcular el **valor de l'estadístic** (z)
6. **Contrastar H_0**. Dues alternatives per fer-ho:
   a. Si $|z| > z_{1-\alpha}$ (unilateral) o $|z| > z_{1-\alpha/2}$ (bilateral) llavors rebutjar H_0 (H_0 és poc versemblant)
   b. Calcular el valor de P → Si $P < \alpha$, llavors rebutjar H_0 (H_0 és poc versemblant)
7. Afegir l'estimació per interval **IC(1-α)**

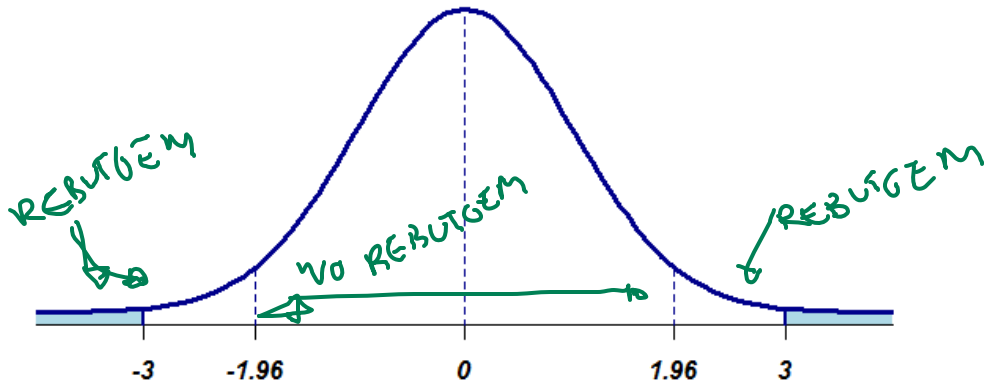
Problema: Com fixem α?

- Valorar el cost que representa una conclusió equivocada
- Definir un llindar arbitrari per fixar què es considera rellevant

Proves d'hipòtesi. Llindars d'acceptació

Es poden trobar els límits d'una regió crítica

- **Bilateral**: a l'esquerra de $z_{\alpha/2}$ (-1.96), i a la dreta de $z_{1-\alpha/2}$ (1.96)
- **Unilateral**:
  - per l'esquerra: per sota de z_{α} (-1.645)
  - per la dreta: per sobre de $z_{1-\alpha}$ (1.645)
- Si l'estadístic cau a la regió crítica (en blau al dibuix), llavors la hipòtesi nul·la (H_0) és dubtosa.
- Si l'estadístic cau a la regió d'acceptació (en blanc al dibuix), llavors no podem rebutjar H_0



Proves d'hipòtesi sobre la μ. Estadístic

- Si σ és coneguda, l'estadístic de referència és:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

- Si σ és desconeguda, l'estadístic de referència és

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

- Sota H_0, tindrà:

  - distribució Normal estàndard, si la grandària mostral és suficient:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim N(0,1) \qquad \text{si} \quad n \geq 100$$

  - distribució t Student amb (n-1) g.ll si la variable estudiada X és Normal

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1} \qquad \text{si} \quad X \sim N$$

Proves d'hipòtesi sobre μ. Exemple

Si en el cas d'una embotelladora de 1 litre tenim: $\bar{x}$ =997, s =10 i n=100. Podem pensar que la mitjana poblacional és μ = 1000 cc?

1. **Variable**: contingut en envasos de 1000cc
2. **Estadístic**: $\hat{t} = (\bar{x} - \mu)/(s/\sqrt{n})$
3. H_0 : μ = 1000cc vs. H_1 : μ ≠ 1000cc
4. **Distribució de l'estadístic** sota H_0 : $(\bar{x} - \mu)/(s\sqrt{n}) \sim N(0,1)$ ja que n = 100;
5. **Càlculs**: $\hat{t} = (\bar{x} - \mu)/(s\sqrt{n}) = \frac{997-1000}{10/\sqrt{100}} = -3$
6. **P-valor** = P(|z| > |-3|) = 0.0027 [**Taules**: 2*(1-0.9987) ; **R**: pnorm(-3)+(1–pnorm(3)]
7. **Conclusió**: com que *P* és menor que α, es rebutja H_0: μ=1000cc

   Conclusió pràctica: ens estan estafant!

8. $\mathbf{IC(\mu, 0.95)} = \bar{x} \mp z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 997 \mp 1.96 \cdot \frac{10}{\sqrt{100}} = [995.04\,, 998.96]$

Nota: per graus llibertat propers a 100, D. t Student és molt propera a DN **R**: funció ***t.test***

Proves d'hipòtesi sobre μ. Exercici

En 9 fitxers, la diferencia D entre els temps d'execució de dos programes de compressió de fitxers ha estat de mitjana 6.71 i desviació 6.00. Acceptant que D ~ N, ¿es pot acceptar que E(D) = μ = 0? (és a dir, acceptar que els dos compressors tarden el mateix en mitjana?)

1. **Variable**: D (diferència en temps)
2. **Estadístic**: $\hat{t} = (\bar{d} - \mu)/(s/\sqrt{n})$
3. **H_0** : $\mu = 0$ vs. $H_1 : \mu \neq 0$
4. **Distribució de l'estadístic sota H0** : $(\bar{d} - \mu)/(s\sqrt{n}) \sim t_{n-1}$ ja que D~N
5. **Càlculs**: $\hat{t} = (\bar{d} - \mu)/(s/\sqrt{n}) = \frac{6.71-0}{6/\sqrt{9}} = 3.355$
6. **P-valor**: $P(|\hat{t}| > |-3.355|) = 0.01$ [R: `pt(-3.355) + (1-pt(3.355)]`
7. **Conclusió**: es rebutja H_0: E(D) = μ = 0 ja que el p-valor (0.01) és < α

   Conclusió pràctica: no triguen el mateix
8. $\mathbf{IC(\mu, 0.95)}$: $\bar{x} \mp t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = 6.71 \mp 2.306 \cdot \frac{6}{\sqrt{9}} = [2.1\ ,\ 11.3]$

R: funció ***t.test***

Proves d'hipòtesi sobre π. Exemple

En el **exemple** anterior del basket, amb n=20 i 8 encerts.

1. **Variable**: resultat de cada tir (canasta o no)
2. **Estadístic**: nombre d'encerts X
3. H_0 : $\pi = 0.80$ (sóc un magnífic tirador) vs. H_1 : $\pi < 0.80$ (no sóc tan bo)
4. Si H_0 es certa: $X \sim B(n, \pi) = B(20, 0.80)$

   Premisses: mostra de tirs independents, amb la mateixa probabilitat d'encert

   [Si n és gran (i π lluny de 0 i de 1), pot ser més simple utilitzar l'aproximació de la Binomial a la Normal (en aquest cas, treballem amb P=X/n enlloc de X): Si H_0 es certa: $P \sim N(\pi, \pi(1-\pi)/n)$. Requereix una premissa addicional: $n \cdot \pi > 5$ i $n \cdot (1-\pi) > 5$ (No s'aplicarà en l'exemple)]
5. **Càlcul de l'estadístic** : x = 8
6. ***P-valor*** = P (X més lluny de H_0 que x) = 0.0001 (*unilateral*)
7. **Conclusió**: **SÍ,** hi ha (forta) evidència en contra de H_0: π=0.80
8. El càlcul de l'**IC** quan la *n* és petita no és senzill, però es pot trobar amb la instrucció de R ***binom.test***. En aquest exemple val [0, 0.61]

Si la mostra és petita obtindreu intervals molt amples... Com ha de ser!

Proves d'hipòtesi sobre π. Exercici

Llencem una moneda 100 vegades i obtenim 63 cares. Està "trucada" la moneda?

1. **Variable**: resultat de cada llançament(cara o creu)
2. **Estadístic**: $Z = (P - \pi)/\sqrt{\pi \cdot (1-\pi)/n}$
3. **H_0 :** $\pi = 0.50$ (moneda equilibrada vs. **H_1 :** $\pi \neq 0.50$ (moneda trucada)
4. **Distribució si H_0 es certa**: $Z \sim N\left(\pi, \sqrt{\pi \cdot (1-\pi)/n}\right) = N(0.5\ , 0.05)$

   Premisses: 1) llançaments independents, amb la mateixa prob. de cara i (2) $n\pi>5$; $n(1-\pi)>5$
5. **Càlcul de l'estadístic**: $= \frac{0.63-0.5}{\sqrt{0.5 \cdot (1-0.5)/100}} = 2.6$
6. **P-valor**: P (X més lluny de H_0 que x) = 0.0094 (*bilateral*)
7. **Conclusió**: **SÍ,** hi ha (certa) evidència en contra de H_0: $\pi = 0.50$
8. **IC95%:**

$$IC(\pi, 0.95) = p \mp z_{0.975} \cdot \sqrt{(\pi \cdot (1-\pi))/n} = 0.63 \mp 1.96 \cdot \sqrt{(0.63 \cdot 0.37)/100} = [0.53, 0.73]$$

Proves d'hipòtesi vs IC

- Per exemple, sobre el valor del paràmetre π en la població:
  - **PH** fa una pregunta "tancada": **¿és $\pi = 0.5$?**
  - **IC** fa una pregunta "oberta": **¿quin es el valor de π ?**
- Donar els resultats sempre amb IC implica:
  - Si es rebutja H_0, dir on es troba el paràmetre
  - Si no es rebutja H_0, quantificar la informació de que es disposa
  - L'IC proporciona informació més fàcil d'interpretar que el P-valor.

Formulari. Proves d'hipòtesi

Paràmetre	Hipòtesi	Estadístic	Premisses	Distribució sota H_0	Criteri Decisió (Risc α)
μ	H : μ = μ_0	$\hat{z} = \frac{\bar{y} - \mu_0}{\sqrt{\sigma^2/n}}$	Y→N ò n≥30 i σ coneguda	$\hat{z} \rightarrow$ N(0,1)	Rebutjar H si $\lvert \hat{z} \rvert > z_{1-\alpha/2}$ ($\lvert \hat{z} \rvert$>1.96 amb α=5%)
μ	H : μ = μ_0	$\hat{t} = \frac{\bar{y} - \mu_0}{\sqrt{s^2/n}}$	Y → N	$\hat{t} \rightarrow t_{n-1}$	Rebutjar H si $\lvert \hat{t} \rvert > t_{n-1,1-\alpha/2}$ ($\lvert \hat{t} \rvert > t_{n-1,0.975}$ amb α=5%)
μ	H : μ = μ_0	$\hat{z} = \frac{\bar{y} - \mu_0}{\sqrt{s^2/n}}$	n ≥ 100	$\hat{z} \rightarrow$ N(0,1)	Rebutjar H si $\lvert \hat{z} \rvert > z_{1-\alpha/2}$ ($\lvert \hat{z} \rvert$>1.96 amb α=5%)
π (Binomial)	H : π = π_0	$\hat{z} = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$	$(1-\pi_0)n \geq 5$ $\pi_0 n \geq 5$	$\hat{z} \rightarrow$ N(0,1)	Rebutjar H si $\lvert \hat{z} \rvert > z_{1-\alpha/2}$ ($\lvert \hat{z} \rvert$>1.96 amb α=5%)
Anexe: λ (Poisson)	H : λ = λ_0	$\hat{z} = \frac{f - \lambda_0}{\sqrt{\lambda_0}}$	$\lambda_0 \geq 5$	$\hat{z} \rightarrow$ N(0,1)	Rebutjar H si $\lvert \hat{z} \rvert > z_{1-\alpha/2}$ ($\lvert \hat{z} \rvert$>1.96 amb α=5%)
σ (normal)	H : σ = σ_0	$\hat{\chi}^2 = \frac{s^2 \cdot (n-1)}{\sigma^2}$	Y → N	$\hat{\chi}^2 \rightarrow \chi^2_{n-1}$	Rebutjar H si $\hat{\chi}^2 < \chi^2_{n-1,\alpha/2}$ o $\hat{\chi}^2 > \chi^2_{n-1,1-\alpha/2}$
En les proves unilaterals s'acumula el valor de P-valor a un sol costat:				H: μ ≤ μ_0 → Rebutjar H si $\hat{z} > z_{1-\alpha}$ H: μ ≥ μ_0 → Rebutjar H si $\hat{z} < -z_{1-\alpha}$	

Annexe: Premissa de Normalitat

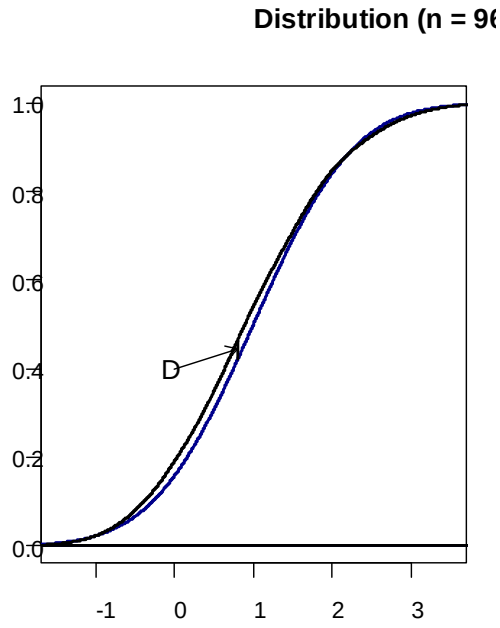
- Hi ha dues mesures que ajuden a valorar el grau d'ajustament, afinitat o similitud a una certa distribució de referència

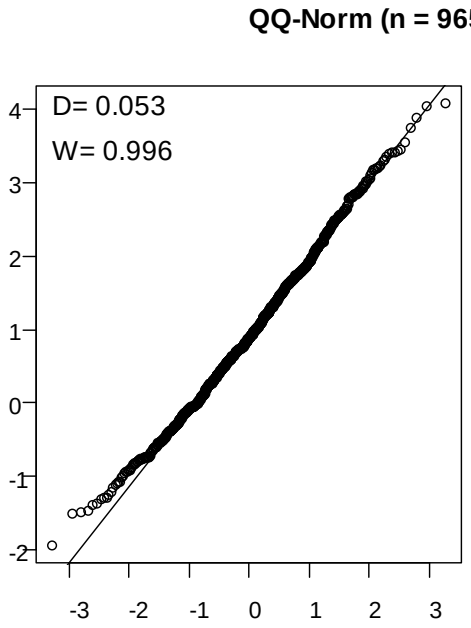
Kolmogorov-Smirnov (Estadístic D)	Shapiro-Wilk (Estadístic W)
Distància màxima entre la funció de distribució empírica i la teòrica. Valors elevats indiquen No Normalitat Entre 0 i 1 (usualment, prop de 0). Valors alts indiquen des-ajustament	Mesura la correlació entre els quantils observats i els teòrics. Valors elevats indiquen Normalitat Entre 0 i 1 (usualment, prop de 1). Valors alts indiquen bon ajustament

- Tots dos estadístics fluctuen a les mostres i han de interpretar-se amb prudència. Proporcionen **P-valors**, però farem només un anàlisi descriptiu i visual (qqplot)
- A continuació mostrem alguns exemples generats a partir de distribucions conegudes i amb diferents mides mostrals.

Annexe: Premissa de Normalitat

- Els paquets solen proporcionar P-valors de la H_0: X ~ N
- No ens hem de fiar massa dels P-valors obtinguts d'aquestes proves
- Es millor emprar les mesures i les eines gràfiques (QQ-Norm, PP-Norm)

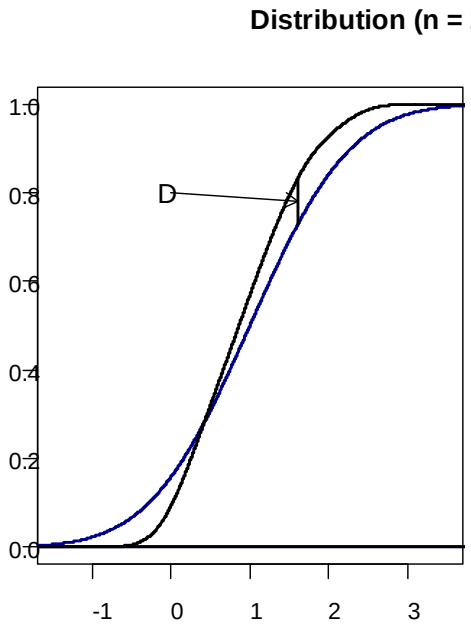


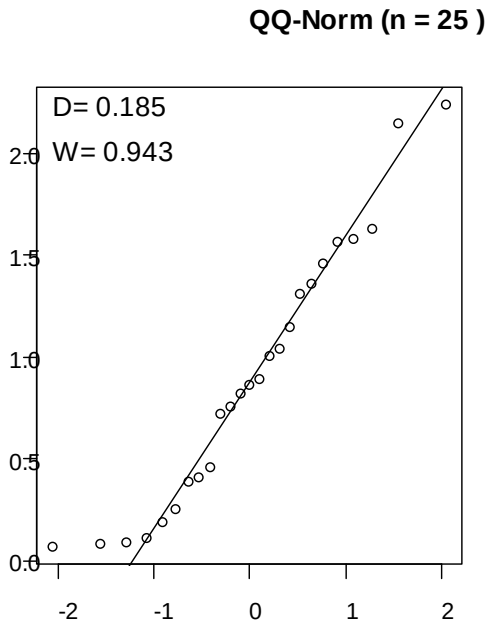


- Aquestes 965 observacions van estar generades seguint una distribució Normal
- D, W i el QQ-Norm mostren que les dades s'ajusten prou bé a una Normal
- No obstant, els p-valors de les proves (0.008 i 0.013) ens farien rebutjar la hipòtesi de Normalitat

Annexe: Premissa de Normalitat

- Aquí tenim un exemple del cas contrari

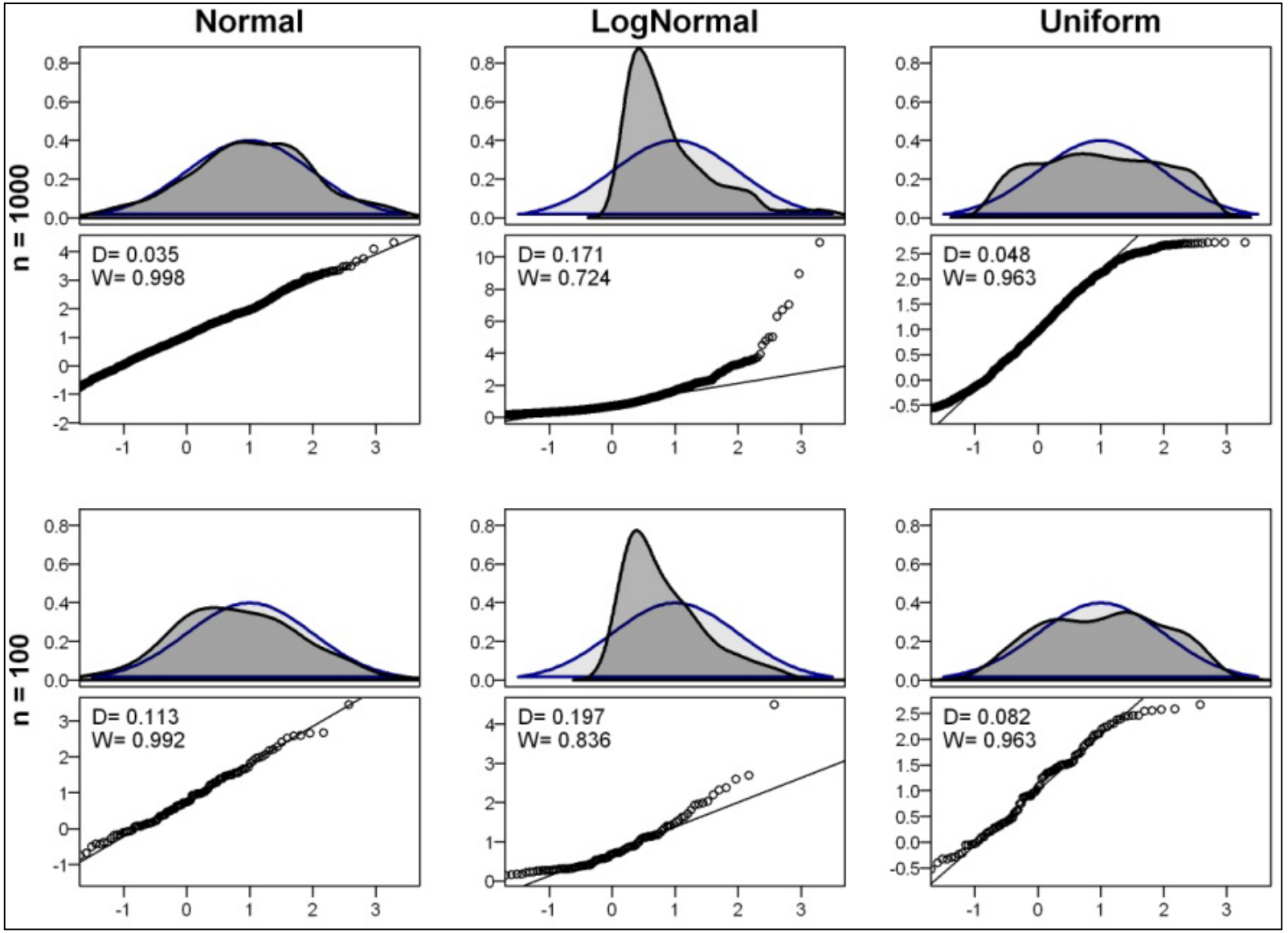




- Aquestes 25 observacions van estar generades seguint una distribució Exponencial
- F_X i D mostren la distància amb la teòrica, encara que W té una bona correlació.
- No obstant, els P-valors de les proves (0.32 i 0.17) no ens fan rebutjar la hipòtesi de Normalitat

Noteu la paradoxa: quan la mostra és petita, és important detectar la No Normalitat, però els P valors fallen. En canvi, quan la mostra és gran, és poc important detectar-la (pel bon comportament asimptòtic dels estimadors) però el P-valor detecta desviacions irrellevants.

Annexe: Premissa de Normalitat (n gran)



Annexe: Premissa de Normalitat (n petita)