Université Internationale
de Casablanca
LAUREATE INTERNATIONAL UNIVERSITIES

Estimation de paramètres d’un modèle probabiliste

Pr. BOUAMAINE A.

Estimation

I. Estimation ponctuelle des paramètres d’une loi de Gauss

1. Estimation de la moyenne

2. Estimation de la variance

II. Intervalle de confiance des paramètres d'une loi de Gauss

1. Estimation de la moyenne

2. Estimation de la variance

III. Estimation de la proportion

1. Estimation ponctuelle
2. Estimation par intervalle

Estimation ponctuelle de la moyenne

Soit X une variable aléatoire réelle qui

suit une loi de Gauss de moyenne m et

d’écart type σ .

Echantillonnage

Pour obtenir des renseignements sur $\theta = m$

On effectue n prélèvements indépendants

$x_1, x_2, \ldots, x_n$

Principe

L'idée de base de l'estimation consiste à chercher une

application ρ définie par :

$$\rho : \begin{array}{ccc} IR^n & \longrightarrow & IR \\ (x_1, x_2, \ldots, x_n) & \longrightarrow & t_n = \rho(x_1, x_2, \ldots, x_n) \end{array}$$

Estimation

la valeur t_n sera prise pour estimation de θ.

Principe

t_n est une réalisation d'une variable aléatoire

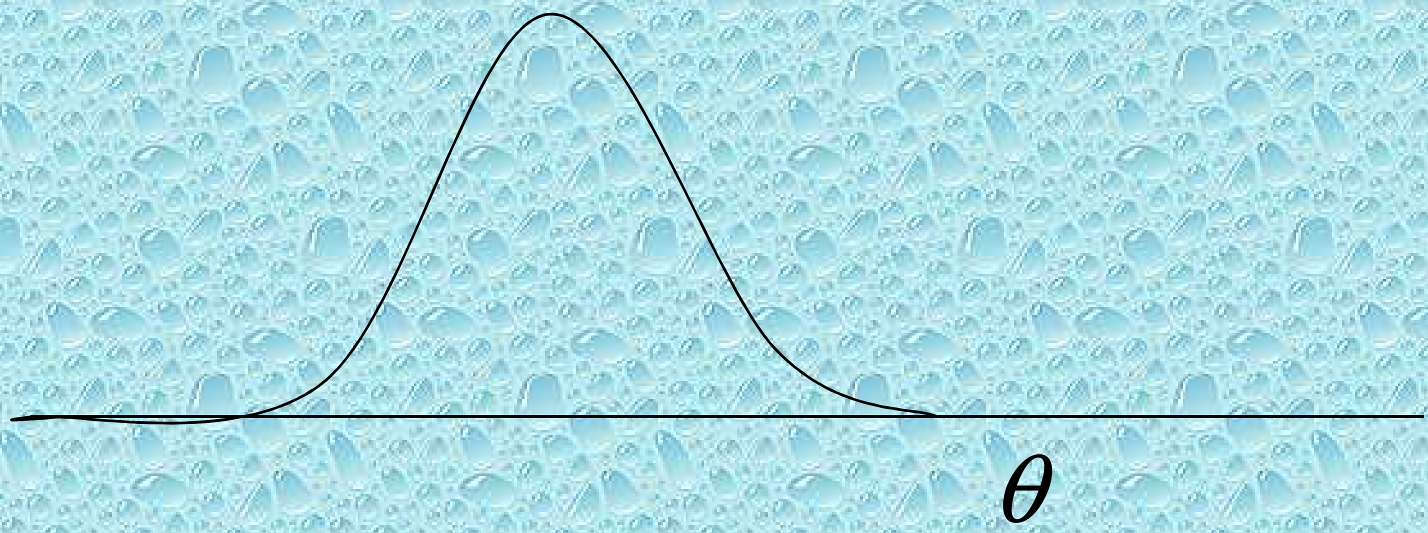
$T_n = \rho (X_1, X_2, \ldots.., X_n)$.

les X_i sont des v.a.r. mutuellement

indépendantes de même loi que X.

Choix de l'application ρ

Supposons que la distribution de T_n vérifie

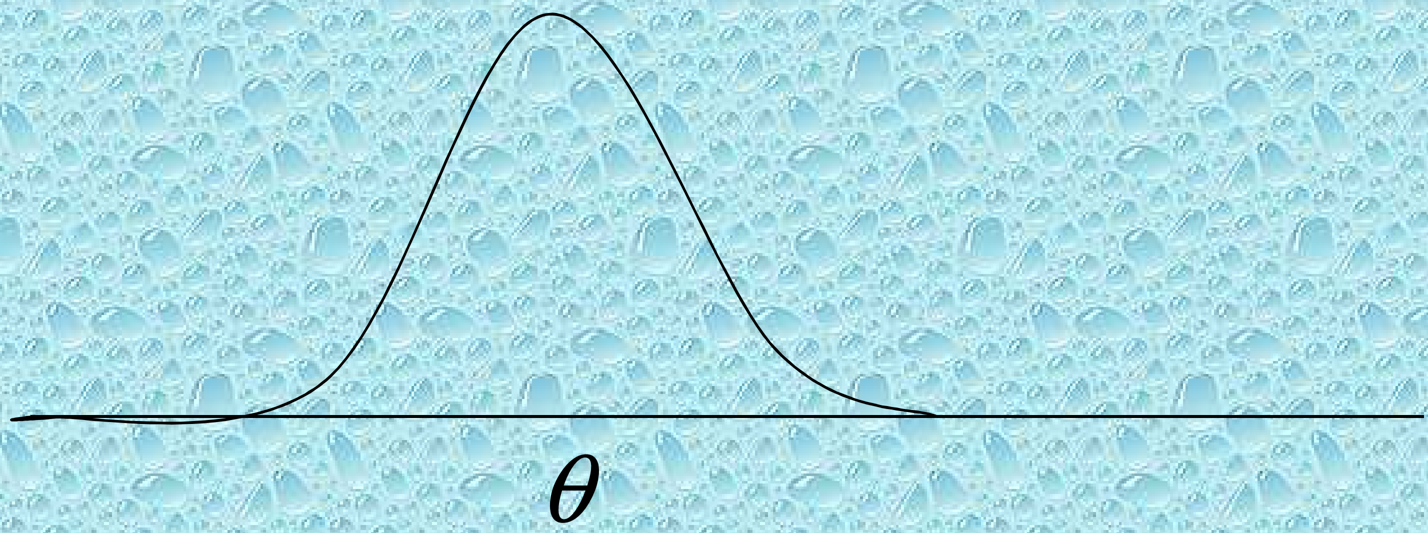


Commentaire

Dans ce cas, aucune réalisation de T_n n'est voisine de θ, d'où T_n ne permet pas d'estimer le paramètre θ.

Choix de l'application ρ

Supposons que la distribution de T_n vérifie



Commentaire

la meilleure estimation de θ est l’observation

la proche de θ, c. à. d. l’intervalle contenant θ

ayant la plus grande probabilité.

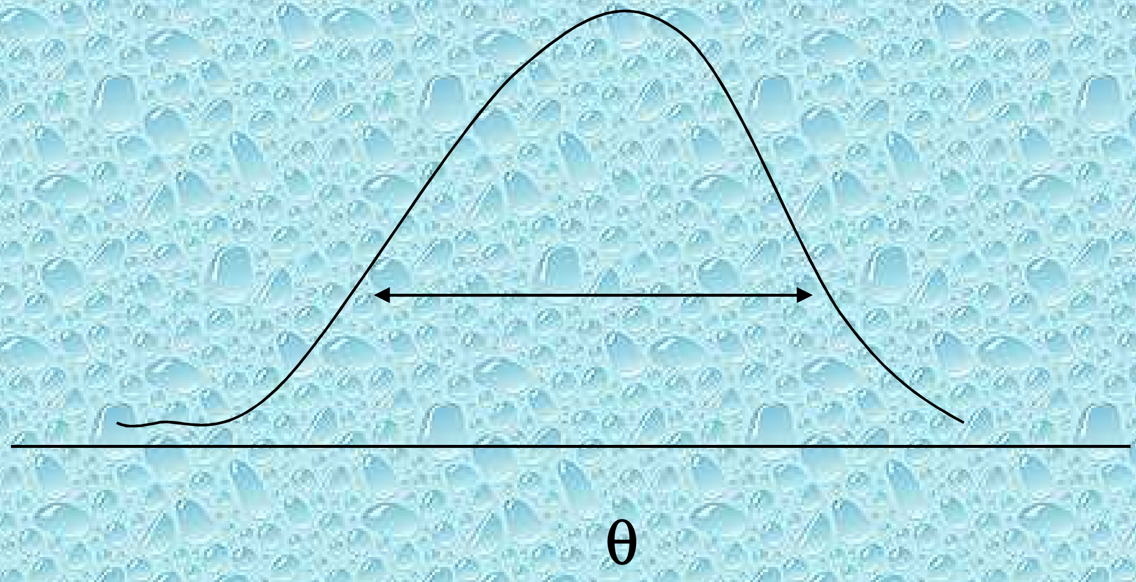
Estimateur non biaisé

T_n est dit un estimateur non biaisé de θ ssi

$$E[\ T_n\ ] = \theta$$

Choix de l’application ρ

- Supposons que la distribution de T_n vérifie

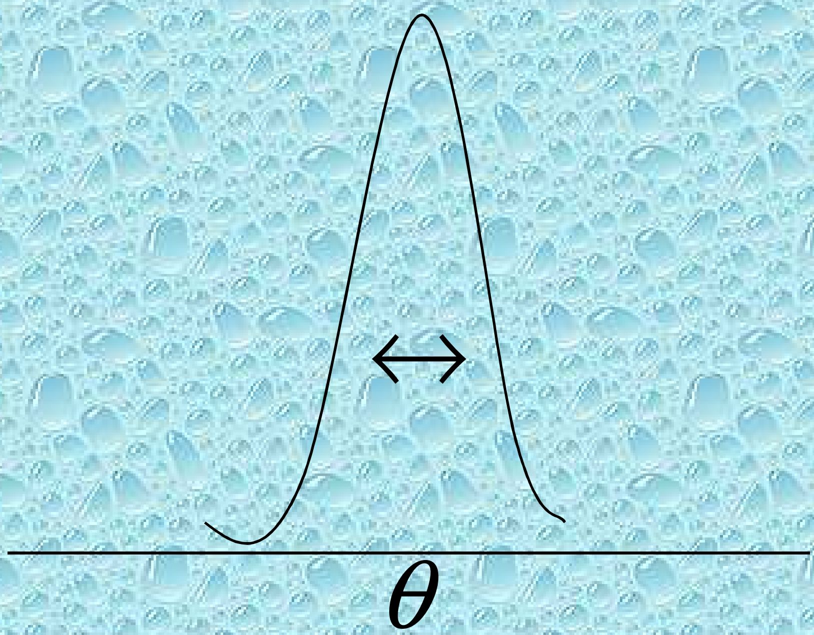


Commentaire

- Cette distribution a une variance grande, on ne peut rien affirmer

Choix de l'application ρ

Supposons que la distribution de T_n vérifie



Commentaire

Une réalisation de cette distribution serait

voisine de θ car sa variance est petite.

Estimateur convergent

Un estimateur convergent est un estimateur

ayant une variance qui tend vers 0

Conséquence

- Il est raisonnable d'imposer à un estimateur T_n de θ quelques qualités, notamment sans biais et convergent.

Variable la moyenne

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Estimateur de la moyenne

$$T_n = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Estimation de la moyenne

$$t_n = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

t_n est une estimation ponctuelle de la moyenne de la loi de Gauss.

Variable variance

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

Propriété

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{X}_\mathrm{i}\right)^2 - \left(\overline{X}\right)^2$$

Estimateur de la variance

$$T_n = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\mathrm{X}_{\mathrm{i}} - \overline{X} \right)^2$$

Estimation ponctuelle de la variance

$$t_n = \frac{n}{n-1} s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\mathrm{x_i} - \overline{x} \right)^2$$

Estimation ponctuelle de l'écart-type

$$\hat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

$\hat{\sigma}_n$ est une estimation ponctuelle de l'écart-type de la loi de Gauss.

Remarque

L’estimation ponctuelle ne donne aucune idée sur l’erreur commise, qui est l’écart entre le paramètre théorique et son estimation.

Intervalle de confiance de la moyenne

Cas 1 : Ecart-type σ est connu

Moyenne aléatoire

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

La variable moyenne suit une loi de Gauss de paramètres m et d’écart-type $\frac{\sigma}{\sqrt{n}}$

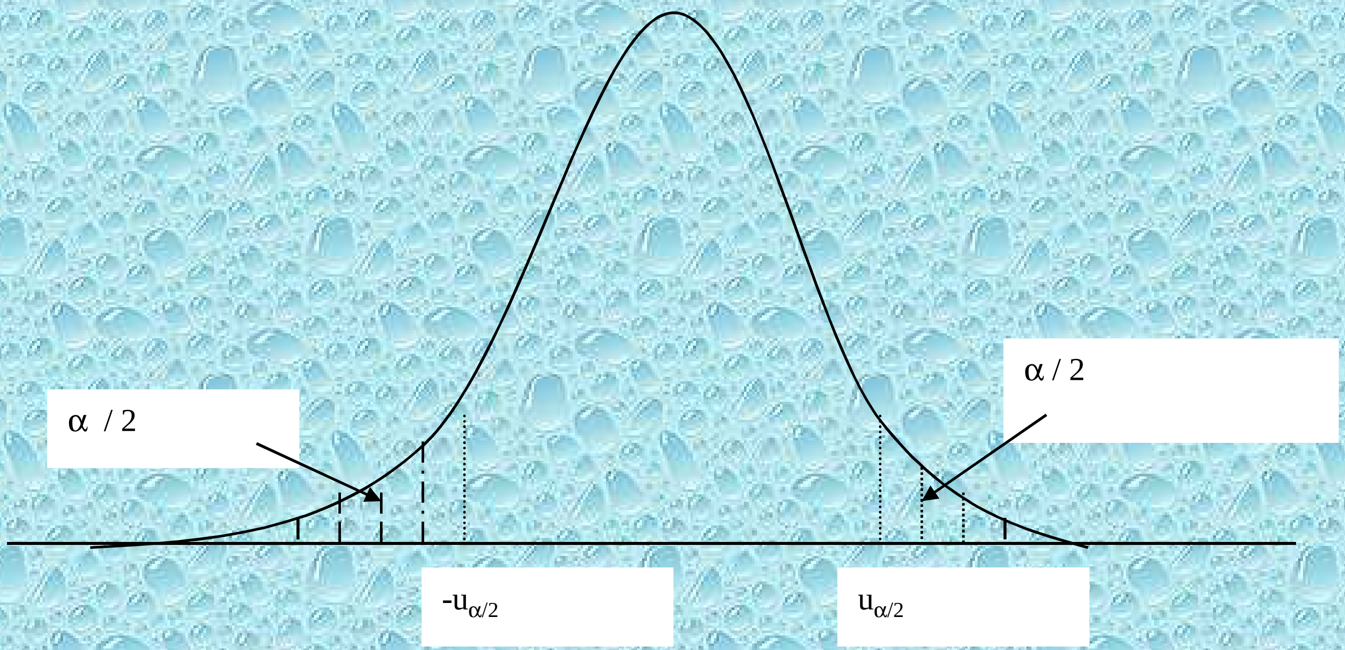
Propriété

On a :

$$L\left(\frac{\overline{X}-\theta}{\frac{\sigma}{\sqrt{n}}}\right)=N(0,1)$$

Propriété

Densité de probabilité de N (0, 1)



On a

$$P\left[\left|\overline{X}-\theta\right| < u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]=1-\alpha$$

On a :

$$P\left[\overline{\mathrm{X}} - u_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \theta < \overline{\mathrm{X}} + u_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$$

Intervalle de confiance aléatoire

$$\left[\overline{X} - u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \ , \ \overline{X} + u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

est appelé intervalle de confiance de la

moyenne de loi de Gauss au seuil α, lorsque

l'écart type est connu

Intervalle de confiance de la moyenne

$$\left[\overline{x} - u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \ , \ \overline{x} + u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

1. n est la taille de l'échantillon
2. σ est l'écart-type de la population supposée connu
3. $\overline{x}$ est la moyenne de l'échantillon

Intervalle de confiance de la moyenne

$$\left[\overline{x} - u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \; , \; \overline{x} + u_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

- α seuil de confiance
- $u_{\alpha/2}$ lu sur la table de N(0,1)

Remarque

On a une probabilité de 1 - α pour que l'intervalle de confiance contienne la moyenne.

Estimation de la moyenne

Cas 2 : Ecart-type σ est inconnu

Propriété

On a :

$$L\left(\frac{\overline{X}-\theta}{\frac{\sigma}{\sqrt{n}}}\right)=N(0,1)$$

Remarque

Comme σ est inconnu, il est raisonnable de

remplacer $\dfrac{\sigma}{\sqrt{n}}$

par un estimateur non biaisé :

Propriété

$\dfrac{S^2}{n-1}$ est un estimateur non biaisé convergent de $\dfrac{\sigma}{\sqrt{n}}$

Posons

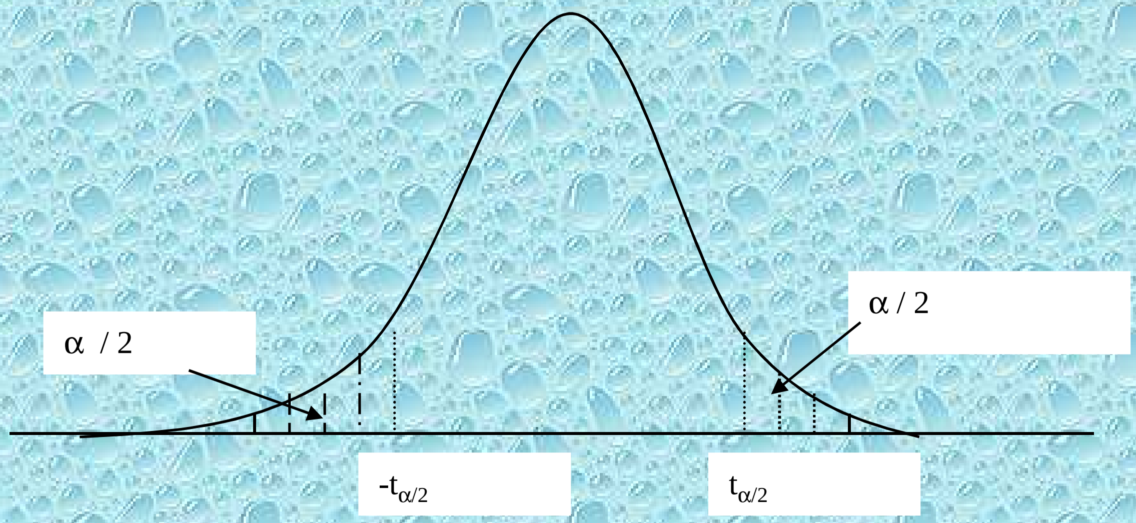
$$T_n = \frac{\overline{X} - \theta}{\frac{S}{\sqrt{n-1}}}$$

Propriété

T_n suit une loi de Student de degré

de libertés n-1

- Densité de probabilité de T_{n-1}



On a :

$$\mathrm{P}\left[\left|\,\overline{X}-\theta\right| < t_{\alpha/2}\,\frac{S}{\sqrt{n-1}}\right]=1-\alpha$$

- **On a :**

$$P\left[\overline{X} - t_{\alpha/2}\frac{S}{\sqrt{n-1}} < \theta < \overline{X} + t_{\alpha/2}\frac{S}{\sqrt{n-1}}\right] = 1-\alpha$$

Intervalle de confiance

L'intervalle aléatoire $\left[\overline{X} -\ t_{\alpha/2}\,\frac{S}{\sqrt{n-1}}\ ;\overline{X} +\ t_{\alpha/2}\,\frac{S}{\sqrt{n-1}}\right]$

est l'intervalle de confiance de la moyenne d'une loi

de Gauss au seuil α lorsque l'écart type est inconnu.

Intervalle de confiance de la moyenne (σ inconnue)

$$\left[\overline{\mathrm{x}} - t_{\alpha/2} \frac{s}{\sqrt{n-1}} \;,\; \overline{\mathrm{x}} + t_{\alpha/2} \frac{s}{\sqrt{n-1}} \right]$$

1. n : la taille de l'échantillon
2. $\overline{x}$: la moyenne de l'échantillon
3. s : écart-type de l'échantillon

Intervalle de confiance de la moyenne

$$\left[\overline{x} - t_{\alpha/2}\frac{s}{\sqrt{n-1}}\ ,\ \overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n-1}}\right]$$

- α seuil de confiance

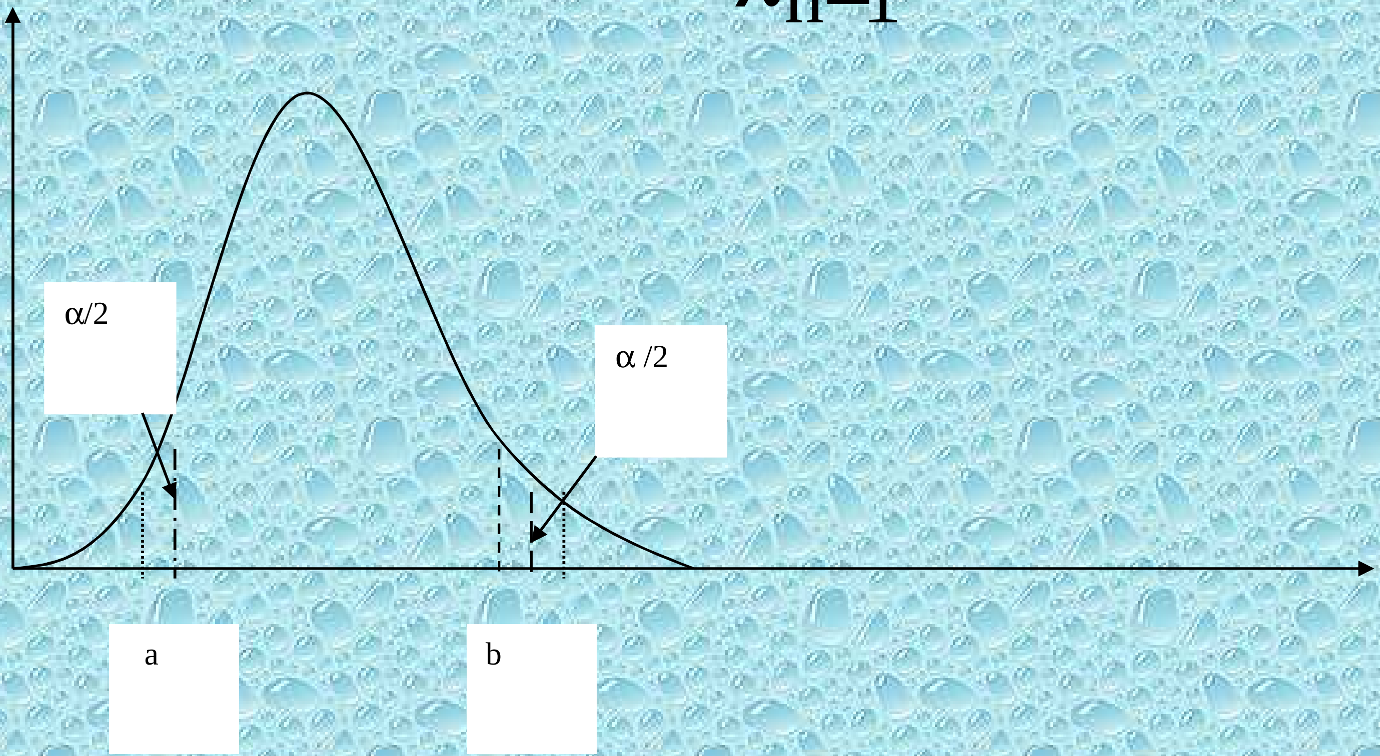
- $t_{\alpha/2}$ lu sur la table de Student de degré de liberté n-1

Intervalle de confiance de l'écart-type

On a :

$$L\left(\frac{nS^2}{\sigma^2}\right) = \chi^2_{n-1}$$

- **Densité de probabilité de** χ^2_{n-1}



Zone d'acceptabilité

On a :

$$P\left[a < \frac{nS^2}{\sigma^2} < b \right] = 1 - \alpha$$

Zone d'acceptabilité

On a :

$$P\left[\frac{nS^2}{b} < \sigma^2 < \frac{nS^2}{a}\right] = 1 - \alpha$$

Intervalle de confiance de la variance

$$\left[\frac{nS^2}{b} \;;\; \frac{nS^2}{a} \right]$$

Intervalle de confiance de la variance

- n : taille de l'échantillon
- α seuil de confiance
- s^2 : variance de l'échantillon
- a, b lus sur la table de khi-deux de degrés de liberté n-1

Estimation d'une proportion

- Soit une population P dont une proportion p d'individus a une propriété donnée A.
- ($0 < p < 1$). On souhaite estimer la proportion inconnue p

Echantillonnage

- On extrait un échantillon au hasard non exhaustif de taille n , k d'entres elles possèdent la propriété A.

Propriété

- k est une réalisation d'une var K_n .

On a : $L(K_n) = B(n, p)$

Estimateur d'une proportion

- Posons :

$$T_n = \frac{K_n}{n}$$

Propriétés

$$E[T_n] = p, \quad Var[T_n] = \frac{p(1-p)}{n}$$

Propriétés

$$E[T_n] = p, \quad Var[T_n] = \frac{p(1-p)}{n}$$

Propriétés

$$E[T_n] = p, \quad Var[T_n] = \frac{p(1-p)}{n}$$

Propriétés

- $T_n = \dfrac{K_n}{n}$ est un estimateur sans biais

et convergent de la proportion p

Estimation ponctuelle d'une proportion

- $t = \hat{p} = \dfrac{k}{n}$ est une estimation ponctuelle de la proportion p d'éléments possédant la propriété A.

Intervalle de confiance d'une proportion

- Sous l'hypothèse que la taille de l'échantillon est suffisamment grande. On a :

$$L\left(\frac{\frac{K_n}{n} - p}{\sqrt{\frac{p(1-p)}{n}}}\right) = N(0,1)$$

Statistique

- Posons :

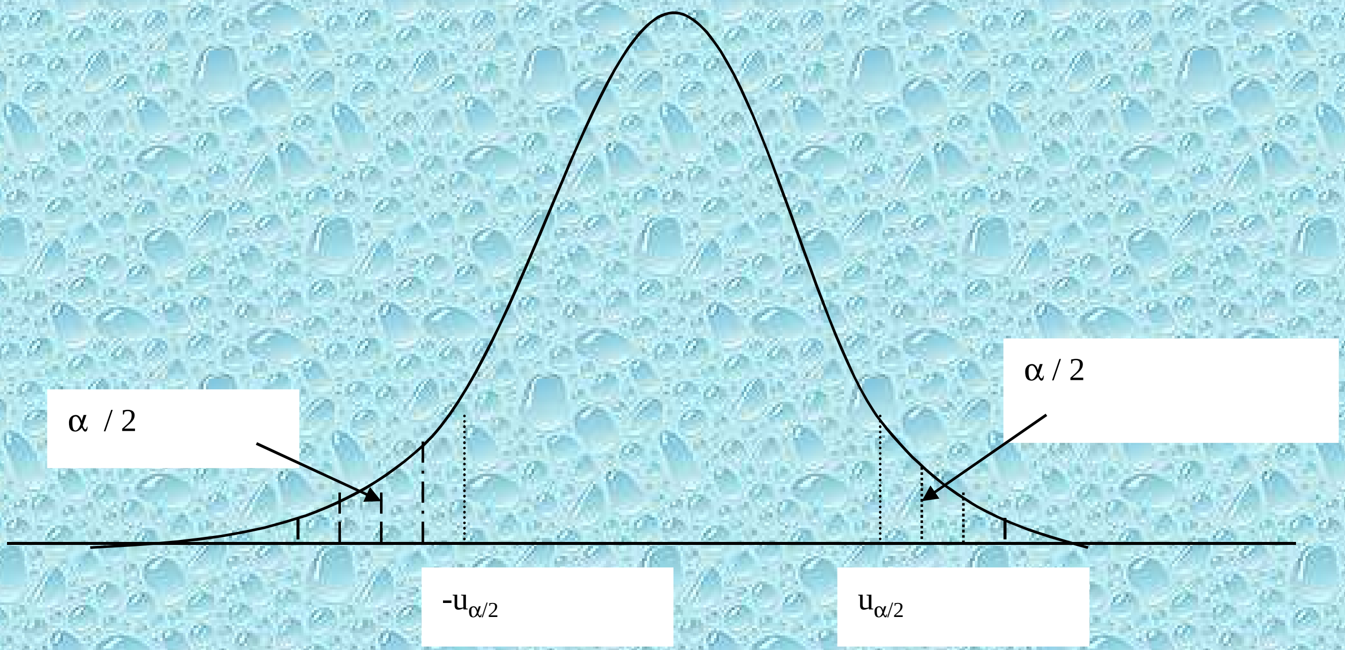
$$T_n = \frac{\frac{K_n}{n} - p}{\sqrt{\frac{K_n(n - K_n)}{n^3}}}$$

Propriété

- La statistique Tn suit une loi de Gauss centré et réduite, sous l'hypothèse que la taille de l'échantillon est importante

Propriété

Densité de probabilité de N (0, 1)



Propriété

- On a :

$$\left[P\left|T_n\right| < u_{\alpha/2}\right] = 1 - \alpha$$

Propriété

$$P\left[\frac{K_n}{n} - u_{\alpha/2}\sqrt{\frac{K_n(1-K_n)}{n^3}} < p < \frac{K_n}{n} + u_{\alpha/2}\sqrt{\frac{K_n(1-K_n)}{n^3}}\right] = 1-\alpha$$

Intervalle de confiance aléatoire d'une proportion

$$\left[\frac{K_n}{n} - u_{\alpha/2}\sqrt{\frac{K_n(1-K_n)}{n^3}}\ ;\ \frac{K_n}{n} + u_{\alpha/2}\sqrt{\frac{K_n(1-K_n)}{n^3}}\right]$$

Intervalle de confiance aléatoire d'une proportion

$$\left[\hat{p} - u_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \; ; \; \hat{p} + u_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

Propriété

- La statistique Tn suit une loi de Gauss centré et réduite, sous l’hypothèse que la taille de l’échantillon est importante

Propriété

- La statistique Tn suit une loi de Gauss centré et réduite, sous l’hypothèse que la taille de l’échantillon est importante

Intervalle de confiance d’une proportion

- n : taille de l’échantillon
- α seuil de confiance
- $\hat{p}$: estimation d’une proportion
- $u_{\alpha/2}$ Lu sur la table de Gauss centré-réduite
-