Intervalles de confiance

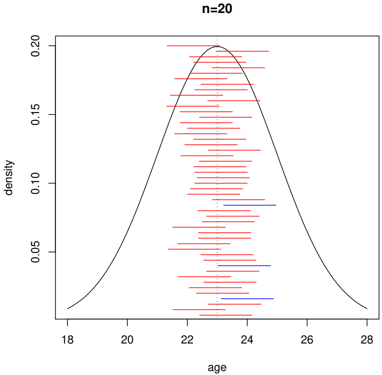
Goual Hafida

UBM Annaba

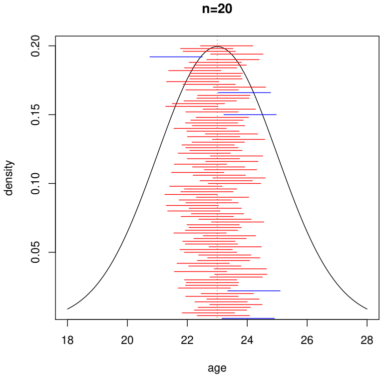
June 5, 2020

Table of contents

Incertitude



Incertitude



Intervalle de confiance de la moyenne

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ alors nous savons que

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{No}(0, 1).$$

Intervalle de confiance de la moyenne

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ alors nous savons que

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim \text{No}(0, 1).$$

Cela signifie que

$$\begin{aligned} \Pr\left(-1.96 < Z < 1.96\right) &= .95. \\ \Pr\left(-1.96 < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < 1.96\right) &= .95. \end{aligned}$$

Intervalle de confiance de la moyenne

Si $X_1, \ldots, X_n \stackrel{iid}{\sim} \mathrm{No}(\mu, \sigma^2)$ alors nous savons que

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathrm{No}(0, 1).$$

Cela signifie que

$$\begin{aligned}
\Pr\left(-1.96 < Z < 1.96\right) &= .95. \\
\Pr\left(-1.96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1.96\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95.
\end{aligned}$$

Intervalle de confiance de la moyenne

Si $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{No}(\mu, \sigma^2)$ alors nous savons que

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathrm{No}(0,1).$$

Cela signifie que

$$\begin{aligned}
\Pr(-1.96 < Z < 1.96) &= .95. \\
\Pr\left(-1.96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1.96\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} - \bar{X} < \mu < -\bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95.
\end{aligned}$$

Intervalle de confiance de la moyenne

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ alors nous savons que

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{No}(0, 1).$$

Cela signifie que

$$\begin{aligned}
\Pr\left(-1.96 < Z < 1.96\right) &= .95. \\
\Pr\left(-1.96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1.96\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} - \bar{X} < \mu < -\bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95. \\
\Pr\left(1.96\frac{\sigma}{\sqrt{n}} + \bar{X} > \mu > \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95.
\end{aligned}$$

Intervalle de confiance de la moyenne

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ alors nous savons que

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{No}(0, 1).$$

Cela signifie que

$$\begin{aligned}
\Pr(-1.96 < Z < 1.96) &= .95. \\
\Pr\left(-1.96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1.96\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95. \\
\Pr\left(-1.96\frac{\sigma}{\sqrt{n}} - \bar{X} < \mu < -\bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95. \\
\Pr\left(1.96\frac{\sigma}{\sqrt{n}} + \bar{X} > \mu > \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95. \\
\Pr\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) &= .95.
\end{aligned}$$

Un intervalle aléatoire

Considérez la quantité

$$\Pr\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = .95,$$

$\bar{X}$ est aléatoire mais μ n'est-ce pas il est fixe.

Un intervalle aléatoire

Considérez la quantité

$$\Pr\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = .95,$$

$\bar{X}$ est aléatoire mais μ n'est-ce pas il est fixe.
L'interprétation de l'équation ci-dessus est un intervalle aléatoire

$$\left(\ell = \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, u = \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right).$$

Un intervalle aléatoire

Considérez la quantité

$$\Pr\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = .95,$$

$\bar{X}$ est aléatoire mais μ n'est-ce pas il est fixe.
L'interprétation de l'équation ci-dessus est un intervalle aléatoire

$$\left(\ell = \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, u = \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right).$$

L'intervalle est centré sur la moyenne de l'échantillon et s'étend dans les deux sens de $1.96\frac{\sigma}{\sqrt{n}}$.

Un intervalle aléatoire

Considérez la quantité

$$\Pr\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = .95,$$

$\bar{X}$ est aléatoire mais μ n'est-ce pas il est fixe.
L'interprétation de l'équation ci-dessus est un intervalle aléatoire

$$\left(\ell = \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, u = \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right).$$

L'intervalle est centré sur la moyenne de l'échantillon et s'étend dans les deux sens de $1.96\frac{\sigma}{\sqrt{n}}$.
Ce que dirait un statisticien
" la probabilité est de .95 que l'intervalle aléatoire inclue la vraie valeur μ."

Définition formelle

Definition

Si $x_1, \ldots, x_n \overset{iid}{\sim} No(\mu, \sigma^2)$ *calculons* $\bar{x}$. *Le* 95% **Intervalle de confiance** *pour* μ *est*

$$\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right),$$

or as $\bar{x} \mp 1.96\frac{\sigma}{\sqrt{n}}$.

Signification d'un IC

Ce que vous voulez qu'un intervalle de confiance dise est
"Ce que vous voulez qu'un intervalle de confiance dise est
$\bar{x} \mp 1.96\frac{\sigma}{\sqrt{n}}$ is .95."

Signification d'un IC

Ce que vous voulez qu'un intervalle de confiance dise est
"Ce que vous voulez qu'un intervalle de confiance dise est
$\bar{x} \mp 1.96\frac{\sigma}{\sqrt{n}}$ is .95."

Signification d'un IC

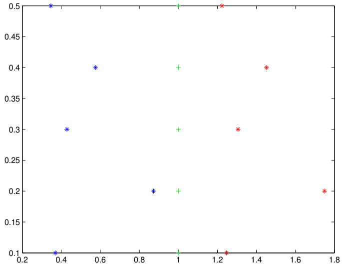
Le 95% IC est interprété comme la limite de la procédure suivante et $\lim_{T \to \infty} val = .05$:

Out $= 0$
pour $t = 1$ à T
    $x_1, \ldots, x_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$
    calculer $\bar{x}$
    Si $\mu \notin \left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$ alors $Out \to Out + 1$
$val = \frac{Out}{T}$

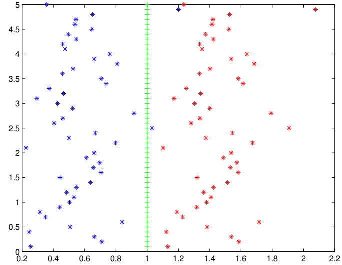
Signification d'un IC

L'IC n'est pas une déclaration concernant l'estimation que vous avez effectuée, mais ce qui se passerait si vous répétiez la même procédure d'estimation encore et encore.

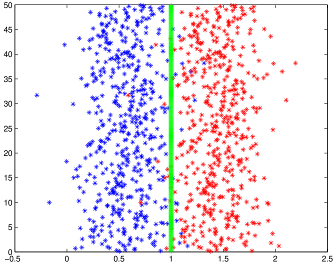
Exemple: $n = 20$ $T = 5$



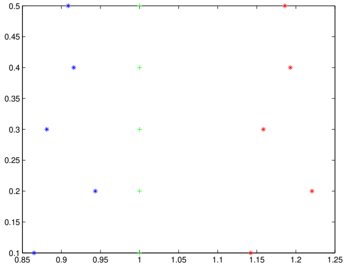
Exemple: $n = 20$ $T = 50$



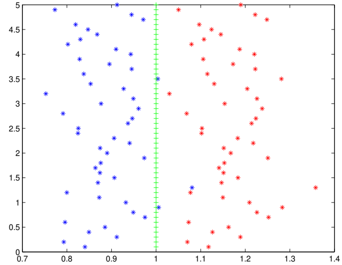
Exemple: $n = 20 \ T = 500$



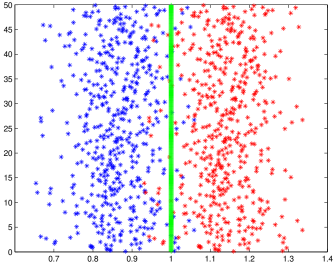
Exemple: $n = 200$ $T = 5$



Exemple: $n = 200$ $T = 50$



Exemple: $n = 200$ $T = 500$



Code

```
T = 500;
n=200;

for i=1:n
x = randn(1,n) + 1;
m = mean(x);
l(1,i) = m - 1.96/sqrt(n);
u(1,i) = m + 1.96/sqrt(n);
end

yv = (1:T)*.1;

plot(l,yv,’b*’);
hold on;
plot(u,yv,’r*’);
plot(1,yv,’g+’);
hold off
```

Niveaux de confiance

Nous pouvons définir tout $100(1-\alpha)\%$ IC pas seulement un 95% IC.

Niveaux de confiance

Nous pouvons définir tout $100(1-\alpha)\%$ IC pas seulement un 95% IC.

Cela se fait en remplaçant 1.96 avec $z_{\alpha/2}$ puisque

$$\Pr\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1 - \alpha.$$

Niveaux de confiance

Nous pouvons définir tout $100(1-\alpha)\%$ IC pas seulement un 95% IC.

Cela se fait en remplaçant 1.96 avec $z_{\alpha/2}$ puisque

$$\Pr\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1-\alpha.$$

Definition

Un $100(1-\alpha)\%$ *IC de* μ *pour une population normale avec connu* σ *est*

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right),$$

ou bien comme $\bar{x} \mp z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Utilisation du TCL

If $X_1, \ldots, X_n$ sont dessinés i.i.d. à partir d'une distribution avec la moyenne μ et la variance σ^2 et n est grande, alors le TCL tient et

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathrm{No}(0, 1).$$

Utilisation du TCL

If $X_1, \ldots, X_n$ sont dessinés i.i.d. à partir d'une distribution avec la moyenne μ et la variance σ^2 et n est grande, alors le TCL tient et

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{No}(0, 1).$$

donc

$$\Pr\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) \approx 1 - \alpha.$$

Utilisation du TCL

If $X_1, \ldots, X_n$ sont dessinés i.i.d. à partir d'une distribution avec la moyenne μ et la variance σ^2 et n est grande, alors le TCL tient et

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim \mathrm{No}(0, 1).$$

donc

$$\mathrm{Pr}\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) \approx 1 - \alpha.$$

Nous ne connaissons presque jamais σ, nous le remplaçons par l'exemple d'écart-type $S = \frac{\sum_i (X_i - \bar{X})^2}{n-1}$ et

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

Utilisation du TCL

If $X_1, \ldots, X_n$ sont dessinés i.i.d. à partir d'une distribution avec la moyenne μ et la variance σ^2 et n est grande, alors le TCL tient et

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim \text{No}(0, 1).$$

donc

$$\Pr\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) \approx 1 - \alpha.$$

Nous ne connaissons presque jamais σ, nous le remplaçons par l'exemple d'écart-type $S = \frac{\sum_i (X_i - \bar{X})^2}{n-1}$ et

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

Maintenant, faites comme si vous étiez dans le cadre normal

Définition formelle

Definition

pour n assez grand ($n > 40$)

$$\bar{x} \mp z_{\alpha/2} \frac{s}{\sqrt{n}}$$

est le grand intervalle de confiance de l'échantillon pour μ avec IC environ $100(1-\alpha)\%$.

Définition formelle

Definition

pour n assez grand $(n > 40)$

$$\bar{x} \mp z_{\alpha/2} \frac{s}{\sqrt{n}}$$

est le grand intervalle de confiance de l'échantillon pour μ *avec IC environ* $100(1-\alpha)\%$.
Cela vaut tant que le TCL est approximativement vrai.

Application 1

Supposons que nous ayons un estimateur $\hat{\theta}$ qui soit

1. normalement distribué
2. approximativement sans biais
3. $\sigma_{\hat{\theta}}$ est disponible.

Application 1

Supposons que nous ayons un estimateur $\hat{\theta}$ qui soit

1. normalement distribué
2. approximativement sans biais
3. $\sigma_{\hat{\theta}}$ est disponible.

Ce qui suit est vrai

$$\mathrm{Pr}\left(-z_{\alpha/2} < \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} < z_{\alpha/2}\right) \approx 1-\alpha$$

Application 1

Supposons que nous ayons un estimateur $\hat{\theta}$ qui soit

1. normalement distribué
2. approximativement sans biais
3. $\sigma_{\hat{\theta}}$ est disponible.

Ce qui suit est vrai

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}} < z_{\alpha/2}\right) \approx 1-\alpha$$

et

$$\hat{\theta} \mp z_{\alpha/2}\frac{s}{\sqrt{n}}$$

est le grand intervalle de confiance de l'échantillon pour θ avec CI environ $100(1-\alpha)\%$.

Application 2: Binomial

Donnons $X \sim \text{Bin}(n, p)$ et $min(np, n(1-p)) \geq 10$ le TCL permet l'approximation normale et $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$.

Alors

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

Application 2: Binomial

Donnons $X \sim \text{Bin}(n, p)$ et $min(np, n(1-p)) \geq 10$ le TCL permet l'approximation normale et $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$.

Alors

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

et nous devons résoudre ce qui précède pour p afin que nous puissions mettre p au milieu.

Application 2: Binomial

Donnons $X \sim \text{Bin}(n, p)$ et $min(np, n(1-p)) \geq 10$ le TCL permet l'approximation normale et $\sigma_{\hat{p}} = \sqrt{p(1-p)/n}$.

Alors

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1 - \alpha$$

et nous devons résoudre ce qui précède pour p afin que nous puissions mettre p au milieu.

Une bonne approximation pour les gros n est

$$\hat{p} \mp z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

est le grand intervalle de confiance de l'échantillon pour μ avec IC environ $100(1-\alpha)\%$.

Binomial

Au lieu de l'approximation

$$\hat{p} \mp z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Nous pouvons essayer de résoudre pour p ce qui suit

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1-\alpha$$

Binomial

Au lieu de l'approximation

$$\hat{p} \mp z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Nous pouvons essayer de résoudre pour p ce qui suit

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1-\alpha$$

alors

$$p = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}}$$

Binomial

Au lieu de l'approximation

$$\hat{p} \mp z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

Nous pouvons essayer de résoudre pour p ce qui suit

$$\Pr\left(-z_{\alpha/2} < \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}\right) \approx 1-\alpha$$

alors

$$p = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1+\frac{z_{\alpha/2}^2}{n}}$$

et

$$\ell = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1+\frac{z_{\alpha/2}^2}{n}}$$

$$u = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{z_{\alpha/2}^2}$$

La distribution t

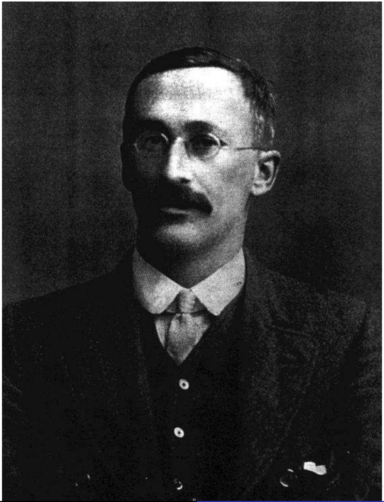
Theorem

Si $\bar{x}$ est la moyenne d'un échantillon aléatoire de taille n tiré d'une distribution normale avec moyenne μ

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

est distribué comme une distribution t avec $\nu = n - 1$ degrés de liberté.

Student: William Sealy Gosset



Propriétés

Soit t_{nu} la densité t avec des degrés de liberté nu

1. t_ν est centré à zéro et en forme de cloche

Propriétés

Soit t_{nu} la densité t avec des degrés de liberté nu

1. t_ν est centré à zéro et en forme de cloche
2. t_ν a une queue plus lourde que la normale

Propriétés

Soit t_{nu} la densité t avec des degrés de liberté nu

1. t_ν est centré à zéro et en forme de cloche
2. t_ν a une queue plus lourde que la normale
3. as ν augmente t_ν a moins de propagation

Propriétés

Soit $t\ _{nu}$ la densité t avec des degrés de liberté *nu*

1. t_ν est centré à zéro et en forme de cloche
2. t_ν a une queue plus lourde que la normale
3. as ν augmente t_ν a moins de propagation
4. as $\lim_{\nu \to \infty} t_\nu \stackrel{dist}{=} \mathrm{No}(0, 1)$ ou bien comme ν augmente t_ν s'approche de la normale standard.

$t_{\alpha,\nu}$ notation

Definition

La notation $t_{\alpha,\nu}$ dénote la valeur z de telle sorte que pour une distribution t avec ν degrés de liberté

$$\Pr(T \geq t_{\alpha,\nu}) = \alpha$$

$t_{\alpha,\nu}$ notation

Definition

La notation $t_{\alpha,\nu}$ dénote la valeur z de telle sorte que pour une distribution t avec ν degrés de liberté

$$\Pr(T \geq t_{\alpha,\nu}) = \alpha$$

ou

$$\Pr(T < t_{\alpha,nu}) = 1 - \alpha.$$

Intervalles de confiance pour les va normaux

Definition

Supposons que $\bar{x}$ et s être la moyenne de l'échantillon et l'écart-type de l'échantillon d'une population normale avec une moyenne μ. Le $100(1-\alpha)\%$ intervalle de confiance pour μ est

$$\bar{x} \mp t_{\alpha/2,\nu} \frac{s}{\sqrt{n}}.$$

Intervalles de confiance pour la variance

Definition

Soit $X_1, .., X_n \overset{iid}{\sim} No(\mu, \sigma^2)$. *Ensuite, la variable aléatoire*

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_i (X_i - \bar{X})^2}{\sigma^2},$$

Intervalles de confiance pour la variance

Definition

Soit $X_1, .., X_n \overset{iid}{\sim} No(\mu, \sigma^2)$. *Ensuite, la variable aléatoire*

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_i (X_i - \bar{X})^2}{\sigma^2},$$

a une distribution chi carré, χ^2_ν, *avec* $\nu = n - 1$ *degrés de liberté.*

χ^2 distribution $\nu = 10$

χ^2 distribution $\nu = 20$

χ^2 distribution $\nu = 30$

χ^2 distribution $\nu = 40$

χ^2 distribution $\nu = 50$

χ^2 distribution $\nu = 60$

χ^2 distribution $\nu = 70$

χ^2 distribution $\nu = 80$

χ^2 distribution $\nu = 90$

χ^2 distribution $\nu = 100$

Valeurs critiques pour χ^2

La distribution χ^2_ν n'est pas symétrique en général. Nous dénotons $\chi^2_{\alpha,\nu}$ comme la valeur telle que $\%100\alpha$ de la zone se trouve à sa droite.

Intervalle de confiance de la variance

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ avec

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Intervalle de confiance de la variance

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ avec

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Cela signifie que

$$\Pr\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}\right) \quad = \quad 1-\alpha.$$

Intervalle de confiance de la variance

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ avec

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Cela signifie que

$$\begin{aligned}
\Pr\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}\right) &= 1-\alpha.\\
\Pr\left(\frac{1}{\chi^2_{1-\alpha/2,n-1}} > \frac{\sigma^2}{(n-1)S^2} > \frac{1}{\chi^2_{\alpha/2,n-1}}\right) &= 1-\alpha.
\end{aligned}$$

Intervalle de confiance de la variance

Si $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{No}(\mu, \sigma^2)$ avec

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Cela signifie que

$$\begin{aligned}
\Pr\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}\right) &= 1-\alpha. \\
\Pr\left(\frac{1}{\chi^2_{1-\alpha/2,n-1}} > \frac{\sigma^2}{(n-1)S^2} > \frac{1}{\chi^2_{\alpha/2,n-1}}\right) &= 1-\alpha. \\
\Pr\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} > \sigma^2 > \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right) &= 1-\alpha.
\end{aligned}$$

Intervalle de confiance de la variance

Si $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{No}(\mu, \sigma^2)$ avec

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Cela signifie que

$$\begin{aligned}
\Pr\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}\right) &= 1-\alpha. \\
\Pr\left(\frac{1}{\chi^2_{1-\alpha/2,n-1}} > \frac{\sigma^2}{(n-1)S^2} > \frac{1}{\chi^2_{\alpha/2,n-1}}\right) &= 1-\alpha. \\
\Pr\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} > \sigma^2 > \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right) &= 1-\alpha. \\
\Pr\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right) &= 1-\alpha.
\end{aligned}$$

Définition formelle

Definition

Soit $x_1, ..., x_n \overset{iid}{\sim} No(\mu, \sigma^2)$ *le* $100(1-\alpha)\%$ **intervalle de confiance** *de* σ^2 *est*

$$\left((n-1)S^2/\chi^2_{\alpha/2,n-1}, (n-1)S^2/\chi^2_{1-\alpha/2,n-1}\right).$$

Limites de confiance

Parfois, nous ne nous soucions que de limiter l'incertitude d'en haut ou d'en bas. Dans ce cas, nous utilisons des limites de confiance.

Limites de confiance

Parfois, nous ne nous soucions que de limiter l'incertitude d'en haut ou d'en bas. Dans ce cas, nous utilisons des limites de confiance. Nous illustrons cela pour la distribution normale avec une variance connue.

Distribution normale variance connue

Si $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{No}(\mu, \sigma^2)$ avec

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathrm{No}(0, 1).$$

Distribution normale variance connue

Si $X_1, \ldots, X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ avec

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{No}(0, 1).$$

cela signifie que

$$\Pr\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > -z_\alpha\right) = 1 - \alpha.$$

Distribution normale variance connue

Si $X_1, ..., X_n \overset{iid}{\sim} \text{No}(\mu, \sigma^2)$ avec

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{No}(0, 1).$$

cela signifie que

$$\begin{aligned} \Pr\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > -z_\alpha\right) &= 1 - \alpha. \\ \Pr\left(\mu < \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right) &= 1 - \alpha. \end{aligned}$$

Définition formelle

Definition

Soit $x_1, ..., x_n \overset{iid}{\sim} No(\mu, \sigma^2)$ *les* $100(1-\alpha)\%$ **intervalles de confiance** *de* μ *sont*

$$\begin{aligned} \mu &< \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}} \\ \mu &> \bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}. \end{aligned}$$

Précision et fiabilité

L'idée derrière un intervalle de confiance est de relier le compromis entre la précision, l'intervalle de confiance et la fiabilité, la confiance ou α.

Précision et fiabilité

L'idée derrière un intervalle de confiance est de relier le compromis entre la précision, l'intervalle de confiance et la fiabilité, la confiance ou α.
Dans le cas normal avec variance connue

$$CI = w = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

and α sont inversement proportionnelles.

Exigences relatives à la taille de l'échantillon

Un problème très courant consiste à trouver la plus petite taille d'échantillon n de telle sorte qu'un niveau particulier de fiabilité et de précision soit satisfait ou étant donné w et *alpha* trouver la plus petite n telle que

$$w = 2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

or

$$n = \left(2z_{\alpha/2}\frac{\sigma}{w}\right)^2.$$