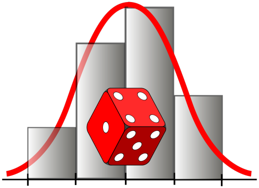
Estimation paramétrique par intervalle de confiance



Module : Techniques d'estimation pour l'ingénieur

Plan

Intervalle de confiance de la variance :

Le problème est le suivant : il faut encadrer σ^2, la variance de la population qui est inconnue. On recherche donc deux valeurs σ_1^2 et σ_2^2 encadrant σ^2 qui vérifie :

$$\mathbb{P}(\sigma_1^2 < \sigma^2 < \sigma_2^2) = 1 - \alpha$$

On distingue deux cas, celui où la moyenne μ est connue et celui où μ est inconnue :

- **Cas moyenne μ connue**

Théorème 5

Un intervalle de confiance de seuil α pour le paramètre σ^2 de la loi $\mathcal{N}(\mu, \sigma^2)$ lorsque μ est connue est :

$$IC(\sigma^2) = \left[\frac{nS_\mu^2}{\chi_{\frac{\alpha}{2},n}}, \frac{nS_\mu^2}{\chi_{1-\frac{\alpha}{2},n}} \right]$$

Démonstration

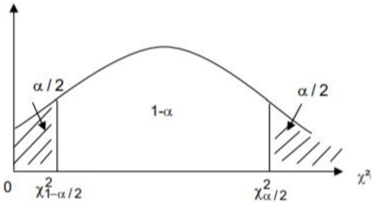
Soit une population normale de variance σ^2 inconnue. La variable aléatoire $\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2}$ suit une loi du $\chi^2(n)$ à $n-$degrés de liberté, qui n'est pas une loi symétrique. L'idée principale de la construction de l'intervalle de confiance I_{σ^2} pour σ^2, avec un risque α fixé, est la suivante On cherche $\chi_{1-\frac{\alpha}{2},n}$ qui vérifie :

$$\mathbb{P}(\chi^2(n) < \chi_{1-\frac{\alpha}{2},n}) = \frac{\alpha}{2}$$

et $\chi_{\frac{\alpha}{2},n}$ qui vérifie :

$$\mathbb{P}(\chi^2(n) > \chi_{\frac{\alpha}{2},n}) = \frac{\alpha}{2}$$

Intervalle de confiance de la variance



Ce qui implique que :

$$\mathbb{P}\left(\chi_{1-\frac{\alpha}{2},n} < \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma^2} < \chi_{\frac{\alpha}{2},n}\right) \quad = \quad 1-\frac{\alpha}{2}-\frac{\alpha}{2} = 1-\alpha$$

alors on montre que

$$\mathbb{P}\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi_{\frac{\alpha}{2},n}} < \sigma^2 < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi_{1-\frac{\alpha}{2},n}}\right) \quad = \quad 1-\alpha$$

Intervalle de confiance de la variance

Si on note $S_\mu^2 := \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{n}$ l'estimateur utilisé pour estimer σ^2 on admet que l'intervalle de confiance de la variance est donné par :

$$IC(\sigma^2) = \left[\frac{nS_\mu^2}{\chi_{\frac{\alpha}{2},n}}, \frac{nS_\mu^2}{\chi_{1-\frac{\alpha}{2},n}} \right]$$

Exemple

Une usine fabrique des câbles. La masse maximale en tonnes supportée par un câble est une variable aléatoire réelle X suivant la loi Normale de moyenne $\mu = 12.2$ et d'écart-type inconnu. Une étude portant sur un échantillon de 20 câbles a donné une variance des charges maximales supportées égales à 2.2 tonnes.

1. Déterminer un intervalle de confiance de σ^2 pour un niveau de confiance 90%.

esprit Se former autrement

Solution

•$n = 20$, $\alpha = 0.1 \longrightarrow \frac{\alpha}{2} = 0.05$.

La masse maximale en tonnes supportée par un câble est une variable aléatoire réelle X suivant la loi Normale de moyenne $\mu = 12.2$ et de variance inconnue mais estimée par $s^2 = 2,2$.

Alors l'intervalle de confiance de la variance est donné par :

$$IC(\sigma^2) = \left[\frac{nS_\mu^2}{\chi_{\frac{\alpha}{2},n}}, \frac{nS_\mu^2}{\chi_{1-\frac{\alpha}{2},n}}\right]$$

A.N :

$$IC(\sigma^2) = \left[\frac{20 \times 2.2}{\chi_{0.05,20}}, \frac{20 \times 2.2}{\chi_{0.95,20}}\right] = \left[\frac{20 \times 2.2}{31.41}, \frac{20 \times 2.2}{10.85}\right]$$
$$= [1.40; 4.05]$$

- **Cas moyenne μ inconnue**

On se propose de donner un intervalle de confiance de niveau de confiance $1-\alpha$ pour σ^2 avec μ inconnue. Dans ce cas l'estimateur ponctuel proposé pour σ^2

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

Par conséquent : $\frac{(n-1)S_n^2}{\sigma^2}$ suit loi du $\chi^2(n-1)$ à $(n-1)-$degrés de liberté. En suivant la même démarche que celle de la situatuion précédente on aboutit au résultat suivant :

Intervalle de confiance de la variance

Théorème 6

Un intervalle de confiance de seuil α pour le paramètre σ^2 de la loi $\mathcal{N}(\mu, \sigma^2)$ lorsque μ est inconnue est :

$$IC(\sigma^2) = \left[\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}} \right]$$

Exercice

On a mesuré la quantité totale d'alcool (exprimée en g/l) contenue dans un échantillon de 10 bouteilles de cidre doux du marché. On a obtenu des valeurs $x_1, x_2, x_3, \ldots, x_{10}$ t.q

$$\sum_{i=1}^{10} x_i = 62 \quad \text{et} \quad \sum_{i=1}^{10} x_i^2 = 388.4124$$

On modélise la quantité d'alcool contenue dans une bouteille, par une variable aléatoire X suivant une loi normale d'espérance μ et de variance σ^2, où les paramètres μ et σ étant inconnus.

1. Proposer des estimations ponctuelles de μ et σ^2 à partir de l'échantillon observé.
2. Construire un intervalle de confiance pour la moyenne au niveau de confiance de $1 - \alpha = 95\%$.
3. Déterminer un intervalle de confiance à 80% de la variance σ^2

Intervalle de confiance de la variance

4. (a) Si n désigne la taille d'un grand échantillon (n $>$ 50), exprimer en fonction de n l'amplitude de l'intervalle de confiance de μ au niveau de confiance de 95%.
   (b) On souhaite construire un intervalle de confiance de μ au niveau de confiance 95% ayant une amplitude de $0,2$ gramme par litre. Quelle est la taille de l'échantillon sachant que $s_n^2 = 0.6$

Solution

1. X : la quantité d'alcool contenue dans une bouteille $\sim \mathcal{N}(\mu, \sigma)$, où les paramètres μ et σ étant inconnus.

$$n = 10 \ , \quad \sum_{i=1}^{10} x_i = 62 \quad \text{et} \quad \sum_{i=1}^{10} x_i^2 = 388.4124$$

• Un estimateur de la moyenne μ est la moyenne empirique donnée par :

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

Alors, une estimation de la moyenne est :

$$\overline{x}_n = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{1}{10}.62 = 6.2$$

Intervalle de confiance de la variance

- Un estimateur ponctuel de la variance est donné par la la variance empirique corrigée :

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2$$

Alors une estimation est donnée par :

$$\begin{aligned}
s_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x_n})^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left[x_i^2 - 2x_i\overline{x_n} + \overline{x_n}^2\right] \\
&= \frac{1}{n-1}\sum_{i=1}^{n}x_i^2 - \frac{2\overline{x_n}}{n-1}\sum_{i=1}^{n}x_i + \frac{n}{n-1}(\overline{x_n})^2
\end{aligned}$$

A.N:

$$s_n^2 = \frac{1}{9} \times 388,4124 - \frac{2 \times 6.2}{9} \times 62 + \frac{10}{9}(6.2)^2 = 0.44$$

2. Construire un intervalle de confiance pour la moyenne au niveau de confiance de $1-\alpha = 95\%$.

   $\bullet n = 10 < 30$, σ^2 inconnue, alors l'intervalle de confiance de la moyenne :

$$IC(\mu) = \left[\overline{X_n} - t_{\frac{\alpha}{2},n-1}\sqrt{\frac{S_n^2}{n}}; \overline{X_n} + t_{\frac{\alpha}{2},n-1}\sqrt{\frac{S_n^2}{n}}\right]$$

$\bullet \alpha = 0.05 \longrightarrow \frac{\alpha}{2} = 0.025 \longrightarrow t_{0.025,9} = 2.262$

Par conséquent :

$$IC(\mu) = [6.2 - 2.262 \times \sqrt{\frac{0.44}{10}}; 6.2 + 2.262 \times \sqrt{\frac{0.44}{10}}] = [5.72; 6.67]$$

3. Déterminer un intervalle de confiance à 80% de la variance σ^2.

   •μ inconnue, alors l'intervalle de confiance de la variance :

$$IC(\sigma^2) = \left[\frac{(n-1)S_n^2}{\chi_{\frac{\alpha}{2},n-1}}, \frac{(n-1)S_n^2}{\chi_{1-\frac{\alpha}{2},n-1}}\right]$$

•$\alpha = 0.2 \longrightarrow \frac{\alpha}{2} = 0.1 \longrightarrow 1 - \frac{\alpha}{2} = 0.9$

Par conséquent :

$$IC(\sigma^2) = \left[\frac{9 \times 0.44}{\chi_{0.1,9}}, \frac{9 \times 0.44}{\chi_{0.9,9}}\right] = \left[\frac{9 \times 0.44}{14.684}, \frac{9 \times 0.44}{4,168}\right]$$

Conclusion :

$$IC(\sigma^2) = [0.26; 0.95]$$

Intervalle de confiance de la variance

4.a Si n désigne la taille d'un grand échantillon (n > 50), exprimer en fonction de n l'amplitude de l'intervalle de confiance de μ au niveau de confiance de 95%.

- Si $n > 50$, et la variance est inconnue, alors:

$$IC(\mu) = \left[\overline{X_n} - z_{\frac{\alpha}{2}}\sqrt{\frac{S_n^2}{n}}; \overline{X_n} + z_{\frac{\alpha}{2}}\sqrt{\frac{S_n^2}{n}}\right]$$

Ce qui implique que

$$\mathcal{A}mp(IC(\mu)) = 2 * z_{\frac{\alpha}{2}}\sqrt{\frac{S_n^2}{n}}$$

Or $\alpha = 0.05 \longrightarrow \frac{\alpha}{2} = 0.025 \longrightarrow z_{\frac{\alpha}{2}} = 1.96$, par conséquent :

$$\mathcal{A}mp(IC(\mu)) = 2 * 1.96\sqrt{\frac{S_n^2}{n}}$$

Intervalle de confiance de la variance

4.b On souhaite construire un intervalle de confiance de μ au niveau de confiance 95% ayant une amplitude de $0,2$ gramme par litre. Quelle est la taille de l'échantillon sachant que $s_n^2 = 0.6$

On a

$$\mathcal{A}mp(IC(\mu)) = 2 * 1.96\sqrt{\frac{S_n^2}{n}} = 0.2$$

Alors

$$n = \frac{(1.96)^2 * 0.6}{(0.1)^2} \approx 230.4$$

On prend $n = 231$.