ACT 2000

Tests

Marie-Pier Côté

École d'actuariat, Université Laval

Hiver 2019

G. Test d'hypothèse (Partie II)

1. Les principes du test d'hypothèse
2. Tester des hypothèses simples
3. Test uniformément le plus puissant
4. Tests pour moyennes et variances
5. Tests courants pour grands échantillons
6. Test du rapport de vraisemblance
7. Tests d'adéquation

Cette matière est basée sur les chapitres 10 et 14.

Où en sommes-nous ?

- Jusqu'à maintenant, nous nous sommes penchés sur l'estimation de paramètre.

Où en sommes-nous ?

- Jusqu'à maintenant, nous nous sommes penchés sur l'estimation de paramètre.
- Les estimateurs ponctuels donnent une seule valeur qui est l'estimation du paramètre inconnu.

Où en sommes-nous ?

- Jusqu'à maintenant, nous nous sommes penchés sur l'estimation de paramètre.
- Les estimateurs ponctuels donnent une seule valeur qui est l'estimation du paramètre inconnu.
- Les estimateurs par intervalle prennent en considération l'incertitude des estimations ponctuelles.

Où en sommes-nous ?

- Jusqu'à maintenant, nous nous sommes penchés sur l'estimation de paramètre.
- Les estimateurs ponctuels donnent une seule valeur qui est l'estimation du paramètre inconnu.
- Les estimateurs par intervalle prennent en considération l'incertitude des estimations ponctuelles.
- Tester des hypothèses statistiques est un problème différent.

Le problème de test d'hypothèse

- On considère un problème statistique dans lequel un paramètre inconnu θ est contenu dans un certain espace des paramètres :

$$\theta \in \Theta.$$

Le problème de test d'hypothèse

- On considère un problème statistique dans lequel un paramètre inconnu θ est contenu dans un certain espace des paramètres :

$$\theta \in \Theta.$$

- On suppose que l'espace des paramètres peut être partitionné en deux sous-ensembles disjoints Θ_0 et Θ_1 tels que

$$\Theta = \Theta_0 \cup \Theta_1.$$

Le problème de test d'hypothèse

- On considère un problème statistique dans lequel un paramètre inconnu θ est contenu dans un certain espace des paramètres :

$$\theta \in \Theta.$$

- On suppose que l'espace des paramètres peut être partitionné en deux sous-ensembles disjoints Θ_0 et Θ_1 tels que

$$\Theta = \Theta_0 \cup \Theta_1.$$

- Plutôt que d'estimer θ, le statisticien doit décider si la valeur inconnue de θ est comprise dans Θ_0 ou dans Θ_1.

Le problème de test d'hypothèse

- On considère un problème statistique dans lequel un paramètre inconnu θ est contenu dans un certain espace des paramètres :

$$\theta \in \Theta.$$

- On suppose que l'espace des paramètres peut être partitionné en deux sous-ensembles disjoints Θ_0 et Θ_1 tels que

$$\Theta = \Theta_0 \cup \Theta_1.$$

- Plutôt que d'estimer θ, le statisticien doit décider si la valeur inconnue de θ est comprise dans Θ_0 ou dans Θ_1.

- Puisque Θ_0 et Θ_1 sont disjoints et que leur union est Θ, il y a seulement deux décisions possibles.

L'hypothèse nulle et la contre-hypothèse

Définition

L'hypothèse $\mathcal{H}_0$ que $\theta \in \Theta_0$ est appelée l'hypothèse nulle :

$$\mathcal{H}_0 : \theta \in \Theta_0.$$

L'hypothèse $\mathcal{H}_1$ que $\theta \in \Theta_1$ est appelée la contre-hypothèse, ou l'hypothèse alternative. En notation,

$$\mathcal{H}_1 : \theta \in \Theta_1.$$

L'hypothèse nulle et la contre-hypothèse

Définition

L'hypothèse $\mathcal{H}_0$ que $\theta \in \Theta_0$ est appelée l'hypothèse nulle :

$$\mathcal{H}_0 : \theta \in \Theta_0.$$

L'hypothèse $\mathcal{H}_1$ que $\theta \in \Theta_1$ est appelée la contre-hypothèse, ou l'hypothèse alternative. En notation,

$$\mathcal{H}_1 : \theta \in \Theta_1.$$

Comme Θ_0 et Θ_1 forment une partition de Θ, exactement une des hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ doit être vraie. Accepter $\mathcal{H}_0$ est donc équivalent à rejeter $\mathcal{H}_1$, et accepter $\mathcal{H}_1$ est équivalent à rejeter $\mathcal{H}_0$.

Exemple

Une politicienne, appelons la Tremblay, est candidate à une élection municipale. Elle gagnera l'élection si elle obtient au moins 50% des votes.

Exemple

Une politicienne, appelons la Tremblay, est candidate à une élection municipale. Elle gagnera l'élection si elle obtient au moins 50% des votes.

Si on imagine que les votes sont un échantillon aléatoire Bernoulli avec probabilité inconnue $p \in (0,1)$, le problème de test est le suivant :

$$\mathcal{H}_0 : \; p \geq 0.5,$$
$$\mathcal{H}_1 : \; p < 0.5.$$

Exemple

Une politicienne, appelons la Tremblay, est candidate à une élection municipale. Elle gagnera l'élection si elle obtient au moins 50% des votes.

Si on imagine que les votes sont un échantillon aléatoire Bernoulli avec probabilité inconnue $p \in (0, 1)$, le problème de test est le suivant :

$$\mathcal{H}_0 : \ p \geq 0.5,$$
$$\mathcal{H}_1 : \ p < 0.5.$$

Autrement dit, l'hypothèse nulle est que Tremblay gagnera l'élection, alors que la contre-hypothèse est qu'elle perdra.

D'autre exemples

- Est-ce que la consommation de carburant de la voiture est plus élevée que celle annoncée par le fabricant ?

D'autre exemples

- Est-ce que la consommation de carburant de la voiture est plus élevée que celle annoncée par le fabricant ?
- Est-ce qu'un nouveau traitement est plus efficace que celui qui est présentement utilisé ?

D'autre exemples

- Est-ce que la consommation de carburant de la voiture est plus élevée que celle annoncée par le fabricant ?
- Est-ce qu'un nouveau traitement est plus efficace que celui qui est présentement utilisé ?
- Est-ce qu'un médicament augmente la pression (ou a d'autres effets secondaires indésirables) ?

D'autre exemples

- Est-ce que la consommation de carburant de la voiture est plus élevée que celle annoncée par le fabricant ?
- Est-ce qu'un nouveau traitement est plus efficace que celui qui est présentement utilisé ?
- Est-ce qu'un médicament augmente la pression (ou a d'autres effets secondaires indésirables) ?
- Est-ce qu'un certain modèle paramétrique est approprié pour les données ?

D'autre exemples

- Est-ce que la consommation de carburant de la voiture est plus élevée que celle annoncée par le fabricant ?
- Est-ce qu'un nouveau traitement est plus efficace que celui qui est présentement utilisé ?
- Est-ce qu'un médicament augmente la pression (ou a d'autres effets secondaires indésirables) ?
- Est-ce qu'un certain modèle paramétrique est approprié pour les données ?

Notez que l'hypothèse nulle et la contre-hypothèse ont des conséquences bien différentes.

Procédure de test

Habituellement, le (la) statisticien(ne) basera sa décision sur les données.

Définition

Une procédure statistique servant à décider s'il faut accepter l'hypothèse nulle ou la contre-hypothèse est appelée un test statistique, ou simplement un test.

Région critique

On considère le problème

$$\mathcal{H}_0 : \theta \in \Theta_0,$$
$$\mathcal{H}_1 : \theta \in \Theta_1,$$

la décision est basée sur un échantillon aléatoire $X_1, \ldots, X_n$ d'une distribution qui dépend du paramètre θ.

Région critique

On considère le problème

$$\mathcal{H}_0 : \theta \in \Theta_0,$$
$$\mathcal{H}_1 : \theta \in \Theta_1,$$

la décision est basée sur un échantillon aléatoire $X_1, \ldots, X_n$ d'une distribution qui dépend du paramètre θ.

Définition

Une procédure de test typique pour le problème ci-haut consiste à créer une partition de l'ensemble $\mathcal{S}$ de tous les résultats possibles pour $\mathbf{X} = (X_1, \ldots, X_n)$ en deux sous-ensembles, $\mathcal{C}$ et $\mathcal{C}^{\complement} = \mathcal{S} \backslash \mathcal{C}$.
On rejettera $\mathcal{H}_0$ si et seulement si $\mathbf{X} \in \mathcal{C}$. Le sous-ensemble $\mathcal{C}$ est appelé la région critique du test.

Fonction de puissance

Définition

On considère une procédure de test avec région critique $\mathcal{C}$. La fonction $\pi : \Theta \to [0,1]$ donnée par

$$\pi(\theta) = \Pr\{(X_1, \ldots, X_n) \in \mathcal{C} | \theta\}$$

est appelée la fonction de puissance du test (*power*). De plus,

$$\sup_{\theta \in \Theta_0} \pi(\theta)$$

est la taille du test (*size*).

Fonction de puissance

Définition

On considère une procédure de test avec région critique $\mathcal{C}$. La fonction $\pi : \Theta \to [0,1]$ donnée par

$$\pi(\theta) = \Pr\{(X_1, \ldots, X_n) \in \mathcal{C} | \theta\}$$

est appelée la fonction de puissance du test (*power*). De plus,

$$\sup_{\theta \in \Theta_0} \pi(\theta)$$

est la taille du test (*size*).

Pour tout $\theta \in \Theta$, la valeur de la fonction de puissance en θ, $\pi(\theta)$, est la probabilité que le test mène au rejet de $\mathcal{H}_0$ si la vraie valeur du paramètre inconnu est θ.

Fonction de puissance idéale

- Idéalement, la fonction de puissance devrait être telle que

$$\pi(\theta) = 0 \quad \Leftrightarrow \quad \theta \in \Theta_0,$$
$$\pi(\theta) = 1 \quad \Leftrightarrow \quad \theta \in \Theta_1.$$

Fonction de puissance idéale

- Idéalement, la fonction de puissance devrait être telle que

$$\begin{aligned} \pi(\theta) = 0 \quad &\Leftrightarrow \quad \theta \in \Theta_0, \\ \pi(\theta) = 1 \quad &\Leftrightarrow \quad \theta \in \Theta_1. \end{aligned}$$

- Un tel test mènerait à la bonne décision avec probabilité 1.

Fonction de puissance idéale

- Idéalement, la fonction de puissance devrait être telle que

$$\begin{aligned} \pi(\theta) = 0 \quad &\Leftrightarrow \quad \theta \in \Theta_0, \\ \pi(\theta) = 1 \quad &\Leftrightarrow \quad \theta \in \Theta_1. \end{aligned}$$

- Un tel test mènerait à la bonne décision avec probabilité 1.
- En pratique, il n'existe pas de test avec la fonction de puissance idéale.

Fonction de puissance idéale

- Idéalement, la fonction de puissance devrait être telle que

$$\begin{aligned} \pi(\theta) = 0 \quad &\Leftrightarrow \quad \theta \in \Theta_0, \\ \pi(\theta) = 1 \quad &\Leftrightarrow \quad \theta \in \Theta_1. \end{aligned}$$

- Un tel test mènerait à la bonne décision avec probabilité 1.
- En pratique, il n'existe pas de test avec la fonction de puissance idéale.
- Par conséquent, un test mènera à une décision incorrecte avec probabilité plus grande que 0.

Exemple

On considère un échantillon $X_1, \ldots, X_n$ de la distribution uniforme $\mathcal{U}(0, \theta)$ avec paramètre inconnu $\theta \in (0, \infty)$.

Exemple

On considère un échantillon $X_1, \ldots, X_n$ de la distribution uniforme $\mathcal{U}(0, \theta)$ avec paramètre inconnu $\theta \in (0, \infty)$.

Nous voulons tester les hypothèses suivantes :

$$\mathcal{H}_0 : \ \theta \in [3, 4],$$
$$\mathcal{H}_1 : \ \theta \in (0, 3) \cup (4, \infty).$$

Exemple

On considère un échantillon $X_1, \ldots, X_n$ de la distribution uniforme $\mathcal{U}(0, \theta)$ avec paramètre inconnu $\theta \in (0, \infty)$.

Nous voulons tester les hypothèses suivantes :

$$\mathcal{H}_0 : \ \theta \in [3, 4],$$
$$\mathcal{H}_1 : \ \theta \in (0, 3) \cup (4, \infty).$$

On rappelle que l'EMV de θ est $M_n = \max(X_1, \ldots, X_n)$ et pour $x \in [0, \theta]$, la fonction de répartition de M_n est $\Pr(M_n \leq x) = (x/\theta)^n$.

Exemple

On considère un échantillon $X_1, \ldots, X_n$ de la distribution uniforme $\mathcal{U}(0, \theta)$ avec paramètre inconnu $\theta \in (0, \infty)$.

Nous voulons tester les hypothèses suivantes :

$$\mathcal{H}_0 : \ \theta \in [3, 4],$$
$$\mathcal{H}_1 : \ \theta \in (0, 3) \cup (4, \infty).$$

On rappelle que l'EMV de θ est $M_n = \max(X_1, \ldots, X_n)$ et pour $x \in [0, \theta]$, la fonction de répartition de M_n est $\Pr(M_n \leq x) = (x/\theta)^n$.

Pour illustrer, on considère un test qui accepte $\mathcal{H}_0$ si et seulement si

$$2.9 \leq \max(X_1, \ldots, X_n) \leq 4.$$

Exemple (suite)

La région critique de ce test est

$$\mathcal{C} = \{(x_1, \dots, x_n) \in (0, \infty)^n \ : \ \max(x_1, \dots, x_n) \in (0, 2.9) \cup (4, \infty)\}.$$

Exemple (suite)

La région critique de ce test est

$$\mathcal{C} = \{(x_1, \ldots, x_n) \in (0, \infty)^n \ : \ \max(x_1, \ldots, x_n) \in (0, 2.9) \cup (4, \infty)\}.$$

La fonction de puissance de ce test est

$$\pi(\theta) = \Pr(M_n < 2.9|\theta) + \Pr(M_n > 4|\theta).$$

Exemple (suite)

La région critique de ce test est

$$\mathcal{C} = \{(x_1, \ldots, x_n) \in (0, \infty)^n \ : \ \max(x_1, \ldots, x_n) \in (0, 2.9) \cup (4, \infty)\}.$$

La fonction de puissance de ce test est

$$\pi(\theta) = \Pr(M_n < 2.9|\theta) + \Pr(M_n > 4|\theta).$$

Si $\theta < 2.9$, alors

$$\Pr(M_n < 2.9|\theta) = 1 \quad \text{et} \quad \Pr(M_n > 4|\theta) = 0$$

de telle sorte que

$$\pi(\theta) = 1.$$

Exemple (suite)

Si $\theta \in [2.9, 4]$, alors

$$\Pr(M_n < 2.9|\theta) = (2.9/\theta)^n, \quad \Pr(M_n > 4|\theta) = 0$$

et donc

$$\pi(\theta) = (2.9/\theta)^n.$$

Exemple (suite)

Si $\theta \in [2.9, 4]$, alors

$$\Pr(M_n < 2.9|\theta) = (2.9/\theta)^n, \quad \Pr(M_n > 4|\theta) = 0$$

et donc

$$\pi(\theta) = (2.9/\theta)^n.$$

Finalement, si $\theta > 4$,

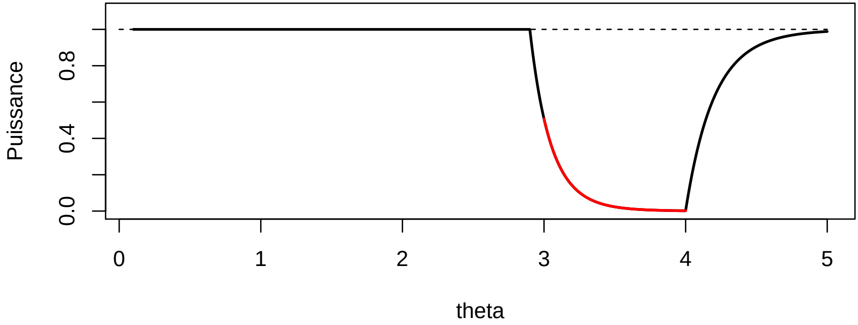
$$\Pr(M_n < 2.9|\theta) = (2.9/\theta)^n, \quad \Pr(M_n > 4|\theta) = 1 - (4/\theta)^n$$

de telle sorte que

$$\pi(\theta) = (2.9/\theta)^n + 1 - (4/\theta)^n.$$

Exemple (suite)

Illustration de la fonction de puissance quand $n = 20$. La ligne rouge montre $\pi(\theta)$ pour $\theta \in \Theta_1 = [3, 4]$. Par exemple, $\pi(3) = 0.5076155$.



Hypothèses simples et composites

On considère un problème

$$\begin{aligned}\mathcal{H}_0 &: \theta \in \Theta_0,\\ \mathcal{H}_1 &: \theta \in \Theta_1.\end{aligned}$$

Hypothèses simples et composites

On considère un problème

$$\mathcal{H}_0 : \theta \in \Theta_0,$$
$$\mathcal{H}_1 : \theta \in \Theta_1.$$

Définition

L'hypothèse $\mathcal{H}_i$ pour $i = 0$ ou $i = 1$ est appelée une hypothèse simple si Θ_i contient une seule valeur de θ. Sinon, l'hypothèse est dite composite.

Hypothèses simples et composites

On considère un problème

$$\mathcal{H}_0 : \theta \in \Theta_0,$$
$$\mathcal{H}_1 : \theta \in \Theta_1.$$

Définition

L'hypothèse $\mathcal{H}_i$ pour $i = 0$ ou $i = 1$ est appelée une hypothèse simple si Θ_i contient une seule valeur de θ. Sinon, l'hypothèse est dite composite.

On note que sous une hypothèse simple, la distribution des observations est complètement spécifiée, alors que sous une hypothèse composite, on sait seulement qu'elle est dans une certaine classe de distributions.

Tester des hypothèses simples

Dans cette section, on suppose que

$$X_1, \ldots, X_n$$

est un échantillon aléatoire d'une distribution avec FMP ou densité $f(\cdot;\theta)$.

Tester des hypothèses simples

Dans cette section, on suppose que

$$X_1, \ldots, X_n$$

est un échantillon aléatoire d'une distribution avec FMP ou densité $f(\cdot;\theta)$.

De plus, on suppose que θ peut prendre seulement deux valeurs :

$$\theta \in \{\theta_0, \theta_1\}.$$

Tester des hypothèses simples

Dans cette section, on suppose que

$$X_1, \ldots, X_n$$

est un échantillon aléatoire d'une distribution avec FMP ou densité $f(\cdot;\theta)$.

De plus, on suppose que θ peut prendre seulement deux valeurs :

$$\theta \in \{\theta_0, \theta_1\}.$$

Les hypothèses simples suivantes sont testées :

$$\mathcal{H}_0 : \ \theta = \theta_0,$$
$$\mathcal{H}_1 : \ \theta = \theta_1.$$

La distribution conjointe de l'échantillon

Sous les postulats précédents, la FMP ou densité conjointe de $(X_1, \dots, X_n)$ est

$$f(\mathbf{x}; \theta_0) = f(x_1; \theta_0) \times \cdots \times f(x_n; \theta_0)$$

si l'hypothèse nulle est vraie,

La distribution conjointe de l'échantillon

Sous les postulats précédents, la FMP ou densité conjointe de $(X_1, \ldots, X_n)$ est

$$f(\mathbf{x}; \theta_0) = f(x_1; \theta_0) \times \cdots \times f(x_n; \theta_0)$$

si l'hypothèse nulle est vraie, ou

$$f(\mathbf{x}; \theta_1) = f(x_1; \theta_1) \times \cdots \times f(x_n; \theta_1)$$

si la contre-hypothèse est vraie.

Deux types d'erreur

Une procédure de test peut faire deux types de décisions incorrectes :

		Décision	
		$\mathcal{H}_0$	$\mathcal{H}_1$
Vérité	$\mathcal{H}_0$	✓	×
	$\mathcal{H}_1$	×	✓

Deux types d'erreur

Une procédure de test peut faire deux types de décisions incorrectes :

		Décision	
		$\mathcal{H}_0$	$\mathcal{H}_1$
Vérité	$\mathcal{H}_0$	✓	×
	$\mathcal{H}_1$	×	✓

Définition 10.1

Soit une procédure de test δ. On fait une erreur de type I si on rejette $\mathcal{H}_0$ alors que $\mathcal{H}_0$ est vraie. On fait une erreur de type II si on accepte $\mathcal{H}_0$ alors $\mathcal{H}_1$ est vraie. Les probabilités d'erreur de type I et II sont notées par

$$\alpha(\delta) = \text{Pr}(\text{rejeter } \mathcal{H}_0 | \theta = \theta_0),$$
$$\beta(\delta) = \text{Pr}(\text{accepter } \mathcal{H}_0 | \theta = \theta_1).$$

Désirable vs. possible

- Si π est la fonction de puissance d'une procédure de test δ,

$$\begin{aligned}\alpha(\delta) &= \Pr(\text{rejeter } \mathcal{H}_0 | \theta = \theta_0) = \pi(\theta_0),\\ \beta(\delta) &= \Pr(\text{accepter } \mathcal{H}_0 | \theta = \theta_1) = 1 - \pi(\theta_1).\end{aligned}$$

Désirable vs. possible

- Si π est la fonction de puissance d'une procédure de test δ,

$$\begin{aligned}\alpha(\delta) &= \Pr(\text{rejeter } \mathcal{H}_0 | \theta = \theta_0) = \pi(\theta_0),\\ \beta(\delta) &= \Pr(\text{accepter } \mathcal{H}_0 | \theta = \theta_1) = 1 - \pi(\theta_1).\end{aligned}$$

- Il est habituellement impossible que $\alpha = \beta = 0$. Il est aussi impossible de trouver une procédure de test pour laquelle α et β sont tous deux très petits.

Désirable vs. possible

- Si π est la fonction de puissance d'une procédure de test δ,

$$\alpha(\delta) = \Pr(\text{rejeter } \mathcal{H}_0 | \theta = \theta_0) = \pi(\theta_0),$$
$$\beta(\delta) = \Pr(\text{accepter } \mathcal{H}_0 | \theta = \theta_1) = 1 - \pi(\theta_1).$$

- Il est habituellement impossible que $\alpha = \beta = 0$. Il est aussi impossible de trouver une procédure de test pour laquelle α et β sont tous deux très petits.

- Il est toutefois possible de trouver un test δ tel que, pour $a, b > 0$,

$$a\alpha(\delta) + b\beta(\delta)$$

est minimisée.

Tests optimaux

Supposons que $a, b > 0$ sont des constantes specifiées.

Théorème

Soit δ^* une procédure de test telle que l'hypothèse $\mathcal{H}_0$ est acceptée si

$$af(\mathbf{x};\theta_0) > bf(\mathbf{x};\theta_1)$$

et $\mathcal{H}_1$ est acceptée si

$$af(\mathbf{x};\theta_0) < bf(\mathbf{x};\theta_1).$$

Si $af(\mathbf{x};\theta_0) = bf(\mathbf{x};\theta_1)$, soit $\mathcal{H}_0$ ou $\mathcal{H}_1$ peut être acceptée.

Tests optimaux

Supposons que $a, b > 0$ sont des constantes specifiées.

Théorème

Soit δ^* une procédure de test telle que l'hypothèse $\mathcal{H}_0$ est acceptée si

$$af(\mathbf{x};\theta_0) > bf(\mathbf{x};\theta_1)$$

et $\mathcal{H}_1$ est acceptée si

$$af(\mathbf{x};\theta_0) < bf(\mathbf{x};\theta_1).$$

Si $af(\mathbf{x};\theta_0) = bf(\mathbf{x};\theta_1)$, soit $\mathcal{H}_0$ ou $\mathcal{H}_1$ peut être acceptée. Alors pour toute autre procédure de test δ,

$$a\alpha(\delta^*) + b\beta(\delta^*) \leq a\alpha(\delta) + b\beta(\delta).$$

Preuve dans le cas discret

Soit $X_1, \ldots, X_n$ un échantillon aléatoire d'une distribution discrète.

Supposons que $\mathcal{C}_\delta$ est la région critique de δ. Alors

$$\begin{aligned}
a\alpha(\delta) + b\beta(\delta) &= a \sum_{\mathbf{x} \in \mathcal{C}_\delta} f(\mathbf{x}; \theta_0) + b \sum_{\mathbf{x} \in \mathcal{C}_\delta^{\complement}} f(\mathbf{x}; \theta_1) \\
&= a \sum_{\mathbf{x} \in \mathcal{C}_\delta} f(\mathbf{x}; \theta_0) + b - b \sum_{\mathbf{x} \in \mathcal{C}_\delta} f(\mathbf{x}; \theta_1) \\
&= b + \sum_{\mathbf{x} \in \mathcal{C}_\delta} \left\{ af(\mathbf{x}; \theta_0) - bf(\mathbf{x}; \theta_1) \right\}.
\end{aligned}$$

Preuve dans le cas discret

Soit $X_1, \ldots, X_n$ un échantillon aléatoire d'une distribution discrète.

Supposons que $\mathcal{C}_\delta$ est la région critique de δ. Alors

$$
\begin{aligned}
a\alpha(\delta) + b\beta(\delta) &= a \sum_{\mathbf{x} \in \mathcal{C}_\delta} f(\mathbf{x}; \theta_0) + b \sum_{\mathbf{x} \in \mathcal{C}_\delta^{\complement}} f(\mathbf{x}; \theta_1) \\
&= a \sum_{\mathbf{x} \in \mathcal{C}_\delta} f(\mathbf{x}; \theta_0) + b - b \sum_{\mathbf{x} \in \mathcal{C}_\delta} f(\mathbf{x}; \theta_1) \\
&= b + \sum_{\mathbf{x} \in \mathcal{C}_\delta} \big\{ af(\mathbf{x}; \theta_0) - bf(\mathbf{x}; \theta_1) \big\}.
\end{aligned}
$$

Clairement, le côté droit est minimisé si

$$
\begin{aligned}
\mathcal{C}_\delta &= \{\mathbf{x} \ : \ af(\mathbf{x}; \theta_0) - bf(\mathbf{x}; \theta_1) < 0\} \\
&= \{\mathbf{x} \ : \ af(\mathbf{x}; \theta_0) < bf(\mathbf{x}; \theta_1)\} = \mathcal{C}_{\delta^*}.
\end{aligned}
$$

Minimiser l'erreur de type I

- Habituellement, la probabilité d'erreur de type I doit être plus petite ou égale à un certain seuil de signification (*significance level*) α. Souvent,

$$\alpha \in \{0.1, 0.05, 0.01\}.$$

Minimiser l'erreur de type I

- Habituellement, la probabilité d'erreur de type I doit être plus petite ou égale à un certain seuil de signification (*significance level*) α. Souvent,

$$\alpha \in \{0.1, 0.05, 0.01\}.$$

- Un test pour lequel la probabilité d'erreur de type I est sous un seuil de signification α est appelé un test au seuil α (*level-α test*).

Minimiser l'erreur de type I

- Habituellement, la probabilité d'erreur de type I doit être plus petite ou égale à un certain seuil de signification (*significance level*) α. Souvent,

$$\alpha \in \{0.1, 0.05, 0.01\}.$$

- Un test pour lequel la probabilité d'erreur de type I est sous un seuil de signification α est appelé un test au seuil α (*level-α test*).

- De plus, alors que l'erreur de type I est plus petite qu'un certain seuil de signification, la probabilité d'une erreur de type II devrait être la plus petite possible.

Lemme de Neyman–Pearson

Théorème 10.1

On suppose que δ^* est une procédure de test qui a la forme suivante pour une constante $k > 0$ donnée : L'hypothèse $\mathcal{H}_0$ est rejetée si

$$f(\mathbf{x}; \theta_0) < kf(\mathbf{x}; \theta_1)$$

et l'hypothèse $\mathcal{H}_0$ est acceptée si $f(\mathbf{x}; \theta_0) > kf(\mathbf{x}; \theta_1)$. Soit $\mathcal{H}_0$ ou $\mathcal{H}_1$ sera acceptée si $f(\mathbf{x}; \theta_0) = kf(\mathbf{x}; \theta_1)$.

Si δ est n'importe quelle autre procédure de test telle que $\alpha(\delta) \leq \alpha(\delta^*)$, alors

$$\beta(\delta) \geq \beta(\delta^*).$$

De plus, si $\alpha(\delta) < \alpha(\delta^*)$, alors $\beta(\delta) > \beta(\delta^*)$.

Preuve

Posons $a = 1$ et $b = k$ dans le théorème précédent.

Preuve

Posons $a = 1$ et $b = k$ dans le théorème précédent.

Alors, pour toute procédure de test δ,

$$\alpha(\delta^*) + k\beta(\delta^*) \leq \alpha(\delta) + k\beta(\delta).$$

Preuve

Posons $a = 1$ et $b = k$ dans le théorème précédent.

Alors, pour toute procédure de test δ,

$$\alpha(\delta^*) + k\beta(\delta^*) \leq \alpha(\delta) + k\beta(\delta).$$

Clairement, si $\alpha(\delta^*) \geq \alpha(\delta)$, alors l'équivalence suivante doit être vérifiée

$$k\beta(\delta^*) \leq k\beta(\delta) \quad \Leftrightarrow \quad \beta(\delta^*) \leq \beta(\delta).$$

Preuve

Posons $a = 1$ et $b = k$ dans le théorème précédent.

Alors, pour toute procédure de test δ,

$$\alpha(\delta^*) + k\beta(\delta^*) \leq \alpha(\delta) + k\beta(\delta).$$

Clairement, si $\alpha(\delta^*) \geq \alpha(\delta)$, alors l'équivalence suivante doit être vérifiée

$$k\beta(\delta^*) \leq k\beta(\delta) \quad \Leftrightarrow \quad \beta(\delta^*) \leq \beta(\delta).$$

On a aussi que si $\alpha(\delta) < \alpha(\delta^*)$, alors $\beta(\delta) > \beta(\delta^*)$.

Interprétation

- Disons qu'on veut que la probabilité d'erreur de type I soit au plus α.

Interprétation

- Disons qu'on veut que la probabilité d'erreur de type I soit au plus α.
- Supposons qu'on trouve une valeur k telle que

$$\alpha(\delta^*) = \Pr\left\{ f(\mathbf{X};\theta_0) < kf(\mathbf{X};\theta_1) \middle| \theta = \theta_0 \right\} = \alpha.$$

Interprétation

- Disons qu'on veut que la probabilité d'erreur de type I soit au plus α.
- Supposons qu'on trouve une valeur k telle que

$$\alpha(\delta^*) = \Pr\left\{ f(\mathbf{X};\theta_0) < kf(\mathbf{X};\theta_1) \middle| \theta = \theta_0 \right\} = \alpha.$$

- Le lemme de Neyman–Pearson garantit que la probabilité d'erreur de type II est la plus petite possible, parmi tous les tests qui ont une probabilité d'erreur de type I plus petite ou égale à α.

Ratio de vraisemblance

Le test au seuil α qui minimise $\beta(\delta)$ rejette $\mathcal{H}_0$ si

$$f(\mathbf{x};\theta_0) < kf(\mathbf{x};\theta_1) \quad \Leftrightarrow \quad \frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} > \frac{1}{k}.$$

Ratio de vraisemblance

Le test au seuil α qui minimise $\beta(\delta)$ rejette $\mathcal{H}_0$ si

$$f(\mathbf{x};\theta_0) < kf(\mathbf{x};\theta_1) \quad \Leftrightarrow \quad \frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} > \frac{1}{k}.$$

Définition

Si $L(\theta)$ dénote la fonction de vraisemblance pour θ basée sur les observations $\mathbf{x} = (x_1, \ldots, x_n)$. La quantité

$$\frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} = \frac{L(\theta_1)}{L(\theta_0)}$$

est le ratio de vraisemblance de l'échantillon.

Exemple 1

Soit $X_1, \ldots, X_n$ un échantillon d'une distribution normale avec $\sigma^2 = 1$ et $\mu \in \{0, 1\}$.

Exemple 1

Soit $X_1, \ldots, X_n$ un échantillon d'une distribution normale avec $\sigma^2 = 1$ et $\mu \in \{0, 1\}$.

Nous voulons tester

$$\begin{aligned} \mathcal{H}_0 &: \mu = 0, \\ \mathcal{H}_1 &: \mu = 1 \end{aligned}$$

au seuil de 1 %.

Exemple 1

Soit $X_1, \ldots, X_n$ un échantillon d'une distribution normale avec $\sigma^2 = 1$ et $\mu \in \{0, 1\}$.

Nous voulons tester

$$\begin{aligned} \mathcal{H}_0 &: \mu = 0, \\ \mathcal{H}_1 &: \mu = 1 \end{aligned}$$

au seuil de 1 %.

Nous voulons donc trouver une procédure de test δ pour laquelle $\alpha(\delta) \leq 0.01$ et $\beta(\delta)$ est la plus petite possible.

Exemple 1 (suite)

Quand $\mathcal{H}_0$ est vraie, l'échantillon est tiré d'une distribution normale centrée réduite et

$$f(\mathbf{x};0) = \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right).$$

Exemple 1 (suite)

Quand $\mathcal{H}_0$ est vraie, l'échantillon est tiré d'une distribution normale centrée réduite et

$$f(\mathbf{x}; 0) = \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right).$$

Quand $\mathcal{H}_1$ est vraie, l'échantillon est tiré de la distribution normale avec moyenne 1 et variance 1. Donc,

$$f(\mathbf{x}; 1) = \frac{1}{(\sqrt{2\pi})^n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - 1)^2\right\}.$$

Exemple 1 (suite)

Le ratio de vraisemblance est

$$\begin{aligned}
\frac{f(\mathbf{x};1)}{f(\mathbf{x};0)} &= \exp\left\{\frac{1}{2}\sum_{i=1}^{n} x_i^2 - \frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2\right\} \\
&= \exp\left\{\frac{1}{2}\sum_{i=1}^{n} x_i^2 - \frac{1}{2}\sum_{i=1}^{n}(x_i^2-2x_i+1)\right\} \\
&= \exp\left(\sum_{i=1}^{n} x_i - \frac{n}{2}\right) \\
&= \exp\left(n\bar{x}_n - \frac{n}{2}\right).
\end{aligned}$$

Exemple 1 (suite)

Selon le lemme de Neyman–Pearson, le test optimal rejette $\mathcal{H}_0$ si, pour un $k > 0$ approprié,

$$\begin{aligned}
\frac{f(\mathbf{x};1)}{f(\mathbf{x};0)} &> \frac{1}{k} && \Leftrightarrow \\
\exp\left(n\bar{x}_n - \frac{n}{2}\right) &> \frac{1}{k} && \Leftrightarrow \\
n\bar{x}_n - \frac{n}{2} &> -\ln(k) && \Leftrightarrow \\
\bar{x}_n &> \underbrace{\frac{1}{2} - \frac{\ln(k)}{n}}_{=k^*}.
\end{aligned}$$

Exemple 1 (suite)

Nous cherchons un test au seuil 0.01, alors on doit trouver k ou k^* tel que

$$\Pr(\bar{X}_n > k^* | \mu = 0) = 0.01$$
$$\Pr(\sqrt{n}\bar{X}_n > \sqrt{n}k^* | \mu = 0) = 0.01.$$

Exemple 1 (suite)

Nous cherchons un test au seuil 0.01, alors on doit trouver k ou k^* tel que

$$\Pr(\bar{X}_n > k^* | \mu = 0) = 0.01$$
$$\Pr(\sqrt{n}\bar{X}_n > \sqrt{n}k^* | \mu = 0) = 0.01.$$

Puisque $\sqrt{n}\bar{X}_n$ suit une $\mathcal{N}(0, 1)$,

$$k^*\sqrt{n} = z_{0.01} = 2.326 \quad \Leftrightarrow \quad k^* = \frac{2.326}{\sqrt{n}}.$$

Exemple 1 (suite)

Nous cherchons un test au seuil 0.01, alors on doit trouver k ou k^* tel que

$$\Pr(\bar{X}_n > k^* | \mu = 0) = 0.01$$
$$\Pr(\sqrt{n}\bar{X}_n > \sqrt{n}k^* | \mu = 0) = 0.01.$$

Puisque $\sqrt{n}\bar{X}_n$ suit une $\mathcal{N}(0, 1)$,

$$k^*\sqrt{n} = z_{0.01} = 2.326 \quad \Leftrightarrow \quad k^* = \frac{2.326}{\sqrt{n}}.$$

Finalement, le test optimal rejette $\mathcal{H}_0$ si

$$\bar{x}_n > \frac{2.326}{\sqrt{n}}.$$

Exemple 1 (suite)

Pour le test optimal au seuil $\alpha = 0.01$, la probabilité d'erreur de type II est

Exemple 1 (suite)

Pour le test optimal au seuil $\alpha = 0.01$, la probabilité d'erreur de type II est

$$\Pr\left(\bar{X}_n \leq \frac{2.326}{\sqrt{n}} \middle| \mu = 1\right) = \Pr\left\{\sqrt{n}\,(\bar{X}_n - 1) \leq 2.326 - \sqrt{n} \middle| \mu = 1\right\}$$
$$= \Phi(2.326 - \sqrt{n}),$$

où Φ est la fonction de répartition de la $\mathcal{N}(0,1)$. Sous $\mathcal{H}_1$, $\sqrt{n}\,(\bar{X}_n - 1)$ suit une normale centrée réduite.

Exemple 1 (suite)

Pour le test optimal au seuil $\alpha = 0.01$, la probabilité d'erreur de type II est

$$\Pr\left(\bar{X}_n \leq \frac{2.326}{\sqrt{n}} \middle| \mu = 1\right) = \Pr\left\{\sqrt{n}\,(\bar{X}_n - 1) \leq 2.326 - \sqrt{n} \middle| \mu = 1\right\}$$
$$= \Phi(2.326 - \sqrt{n}),$$

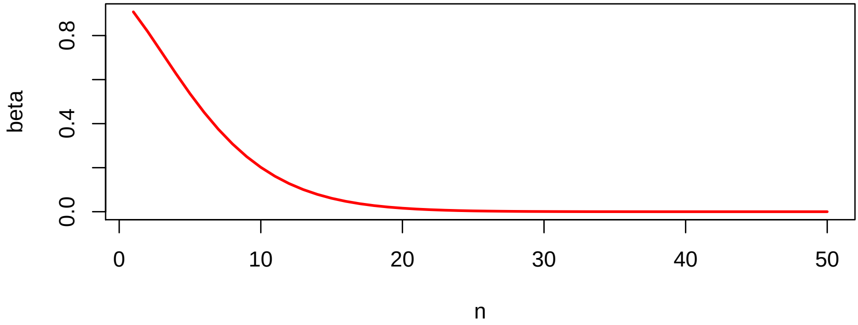
où Φ est la fonction de répartition de la $\mathcal{N}(0,1)$. Sous $\mathcal{H}_1$, $\sqrt{n}\,(\bar{X}_n - 1)$ suit une normale centrée réduite.

Contrairement à la probabilité d'erreur de type I, qui est fixée à 0.01, la probabilité d'erreur de type II dépend de la taille d'échantillon n. Par exemple, quand $n = 20$,

$$\beta = 0.01594.$$

Exemple 1 (la fin !)

Le graphique montre la probabilité d'erreur de type II en fonction de la taille d'échantillon n.



Choisir un seuil de signification

- Le seuil de signification α choisi réflète à quel point les conséquences d'une erreur de type I sont sérieuses.

Choisir un seuil de signification

- Le seuil de signification α choisi réflète à quel point les conséquences d'une erreur de type I sont sérieuses.
- Habituellement,
$$\alpha \in \{0.1, 0.05, 0.01\}.$$
Les valeurs $\alpha = 0.1$ et $\alpha = 0.01$ sont choisies lorsque les conséquences d'une erreur de type I sont faible et sérieuses, respectivement.

Choisir un seuil de signification

- Le seuil de signification α choisi réflète à quel point les conséquences d'une erreur de type I sont sérieuses.

- Habituellement,
$$\alpha \in \{0.1, 0.05, 0.01\}.$$
Les valeurs $\alpha = 0.1$ et $\alpha = 0.01$ sont choisies lorsque les conséquences d'une erreur de type I sont faible et sérieuses, respectivement.

- Les chercheurs utilisent souvent la valeur $\alpha = 0.05$.

Un regard critique

Dans l'**Exemple 1**, le test optimal au seuil $\alpha = 0.01$ rejette $\mathcal{H}_0$ si

$$\bar{x}_n > \frac{z_\alpha}{\sqrt{n}} = \frac{2.326}{\sqrt{n}}$$

et la probabilité d'erreur de type II est

$$\Phi(z_\alpha - \sqrt{n}) = \Phi(2.326 - \sqrt{n}).$$

Un regard critique

Dans l'**Exemple 1**, le test optimal au seuil $\alpha = 0.01$ rejette $\mathcal{H}_0$ si

$$\bar{x}_n > \frac{z_\alpha}{\sqrt{n}} = \frac{2.326}{\sqrt{n}}$$

et la probabilité d'erreur de type II est

$$\Phi(z_\alpha - \sqrt{n}) = \Phi(2.326 - \sqrt{n}).$$

Regardez ce qui arrive pour différentes tailles d'échantillon !

n	$\alpha(\delta^*)$	$\beta(\delta^*)$	k^*	k
1	0.01	0.91	2.33	0.16
25	0.01	0.0038	0.47	2.38
100	0.01	8×10^{-15}	0.23	4.09×10^{11}

Un regard critique (suite)

D'un autre côté, si l'erreur de type I est vue comme beaucoup plus sérieuse qu'une erreur de type II, et qu'on utilise un test qui minimise

$$100\alpha(\delta) + \beta(\delta),$$

alors ce test rejette $\mathcal{H}_0$ si et seulement si

$$\frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} = \exp\left(n\bar{x}_n - \frac{n}{2}\right) > 100 \quad \Leftrightarrow \quad \bar{x}_n > \frac{1}{2} + \frac{\ln(100)}{n}.$$

Un regard critique (suite)

D'un autre côté, si l'erreur de type I est vue comme beaucoup plus sérieuse qu'une erreur de type II, et qu'on utilise un test qui minimise

$$100\alpha(\delta) + \beta(\delta),$$

alors ce test rejette $\mathcal{H}_0$ si et seulement si

$$\frac{f(\mathbf{x};\theta_1)}{f(\mathbf{x};\theta_0)} = \exp\left(n\bar{x}_n - \frac{n}{2}\right) > 100 \quad \Leftrightarrow \quad \bar{x}_n > \frac{1}{2} + \frac{\ln(100)}{n}.$$

n	$\alpha(\delta)$	$\beta(\delta)$	k^*	k
1	1.7×10^{-7}	1	5.11	0.01
25	3.1×10^{-4}	0.057	0.68	0.01
100	2.4×10^{-8}	2.8×10^{-6}	0.55	0.01

Conclusions

- Un seuil de signification pré-spécifié est subjectif.

Conclusions

- Un seuil de signification pré-spécifié est subjectif.
- Un seuil de signification donné n'est pas nécessairement approprié pour toutes les tailles d'échantillon. Par exemple, un seuil de signification qui est approprié pour un petit échantillon ne l'est peut-être pas pour un grand n. Il est peut-être trop grand, et le test serait plus prudent envers les erreurs de type II que celles de type I.

Conclusions

- Un seuil de signification pré-spécifié est subjectif.
- Un seuil de signification donné n'est pas nécessairement approprié pour toutes les tailles d'échantillon. Par exemple, un seuil de signification qui est approprié pour un petit échantillon ne l'est peut-être pas pour un grand n. Il est peut-être trop grand, et le test serait plus prudent envers les erreurs de type II que celles de type I.
- Si on communique seulement si $\mathcal{H}_0$ est rejetée ou pas, notre décision est basée sur le seuil de signification, qui est peut-être inapproprié.

Statistique de test

La plupart des tests sont basés sur une statistique de test

$$T(X_1, \ldots, X_n).$$

Statistique de test

La plupart des tests sont basés sur une statistique de test

$$T(X_1, \ldots, X_n).$$

La région critique du test au seuil α est souvent de la forme

$$\mathcal{C} = \{(x_1, \ldots, x_n) \; : \; T(x_1, \ldots, x_n) \geq c_\alpha\}.$$

Statistique de test

La plupart des tests sont basés sur une statistique de test

$$T(X_1, \ldots, X_n).$$

La région critique du test au seuil α est souvent de la forme

$$\mathcal{C} = \{(x_1, \ldots, x_n) \ : \ T(x_1, \ldots, x_n) \geq c_\alpha\}.$$

Dans ce contexte, c_α est aussi appelée la valeur critique du test.

Exemples

Dans l'**exemple 1**, le test optimal au seuil α rejette $\mathcal{H}_0$ si

$$\bar{X}_n \geq k^*$$

pour une valeur critique k^* choisie correctement.
La statistique du test est $\bar{X}_n$.

Exemples

Dans l'**exemple 1**, le test optimal au seuil α rejette $\mathcal{H}_0$ si

$$\bar{X}_n \geq k^*$$

pour une valeur critique k^* choisie correctement.
La statistique du test est $\bar{X}_n$.

On observe qu'une grande valeur de la statistique du test est moins vraisemblable sous $\mathcal{H}_0$.

Seuil observé ou valeur p

Définition 10.2

Pour une statistique de test T donnée, le seuil observé ou la valeur p (*attained significance level, or p-value*) est le plus petit seuil de signification pour lequel $\mathcal{H}_0$ serait rejetée sur la base des données observées.

Seuil observé ou valeur *p*

Définition 10.2

Pour une statistique de test T donnée, le seuil observé ou la valeur p (*attained significance level, or p-value*) est le plus petit seuil de signification pour lequel $\mathcal{H}_0$ serait rejetée sur la base des données observées.

Autrement dit, le seuil observé est la probabilité sous $\mathcal{H}_0$ que la statistique du test T prenne des valeurs qui sont au moins aussi improbable sous $\mathcal{H}_0$ que la valeur observée t de T.

Seuil observé ou valeur *p*

Définition 10.2

Pour une statistique de test T donnée, le seuil observé ou la valeur p (*attained significance level, or p-value*) est le plus petit seuil de signification pour lequel $\mathcal{H}_0$ serait rejetée sur la base des données observées.

Autrement dit, le seuil observé est la probabilité sous $\mathcal{H}_0$ que la statistique du test T prenne des valeurs qui sont au moins aussi improbable sous $\mathcal{H}_0$ que la valeur observée t de T.

- **$\mathcal{H}_0$ est rejetée au seuil α si la valeur p est plus petite ou égale à α.**

Seuil observé ou valeur *p*

Définition 10.2

Pour une statistique de test T donnée, le seuil observé ou la valeur p (*attained significance level, or p-value*) est le plus petit seuil de signification pour lequel $\mathcal{H}_0$ serait rejetée sur la base des données observées.

Autrement dit, le seuil observé est la probabilité sous $\mathcal{H}_0$ que la statistique du test T prenne des valeurs qui sont au moins aussi improbable sous $\mathcal{H}_0$ que la valeur observée t de T.

- **$\mathcal{H}_0$ est rejetée au seuil α si la valeur p est plus petite ou égale à α.**
- **Le seuil observé n'est PAS la probabilité que $\mathcal{H}_0$ soit vraie.**

Exemple 1 (suite)

Dans l'**exemple 1**, le seuil observé est

$$\Pr(\bar{X}_n \geq \bar{x}_n | \theta = \theta_0),$$

i.e, c'est le plus petit seuil de signification auquel on peut rejeter $\mathcal{H}_0$.

Exemple 1 (suite)

Dans l'**exemple 1**, le seuil observé est

$$\Pr(\bar{X}_n \geq \bar{x}_n | \theta = \theta_0),$$

i.e, c'est le plus petit seuil de signification auquel on peut rejeter $\mathcal{H}_0$.

On trouve

$$\Pr(\bar{X}_n \geq \bar{x}_n | \theta = \theta_0) = \Pr(\sqrt{n}\bar{X}_n \geq \sqrt{n}\bar{x}_n | \theta = \theta_0) = 1 - \Phi(\sqrt{n}\bar{x}_n).$$

Par exemple, si $n = 10$ et $\bar{x}_n = 0.56$,

```
1-pnorm(sqrt(10)*0.56)

## [1] 0.0382907
```

Remarque finale TRÈS importante

- Si le test rejette $\mathcal{H}_0$ au seuil de signification α, alors, la probabilité d'erreur de type I est au plus égale à α. Donc, il est considéré **prudent de rejeter l'hypothèse nulle**.

Remarque finale TRÈS importante

- Si le test rejette $\mathcal{H}_0$ au seuil de signification α, alors, la probabilité d'erreur de type I est au plus égale à α. Donc, il est considéré **prudent de rejeter l'hypothèse nulle**.

- Si le test ne rejette pas $\mathcal{H}_0$ au seuil de signification α, on a aucune idée de l'ampleur de la probabilité d'erreur de type II. Donc, il n'est **pas prudent d'accepter l' hypothèse nulle**.

Remarque finale TRÈS importante

- Si le test rejette $\mathcal{H}_0$ au seuil de signification α, alors, la probabilité d'erreur de type I est au plus égale à α. Donc, il est considéré **prudent de rejeter l'hypothèse nulle**.

- Si le test ne rejette pas $\mathcal{H}_0$ au seuil de signification α, on a aucune idée de l'ampleur de la probabilité d'erreur de type II. Donc, il n'est **pas prudent d'accepter l' hypothèse nulle**. On dit plutôt:

Il n'y a pas assez d'information dans les données pour rejeter l'hypothèse nulle au seuil de signification choisi.

ou simplement que $\mathcal{H}_0$ n'est pas rejetée.

Problèmes plus complexes

- On suppose que $X_1, \ldots, X_n$ est un échantillon aléatoire d'une distribution avec $f(x;\theta)$, où le paramètre $\theta \in \Theta \subseteq \mathbb{R}$ est inconnu. Comme précédemment, on veut tester

$$\mathcal{H}_0 \; : \; \theta \in \Theta_0,$$
$$\mathcal{H}_1 \; : \; \theta \in \Theta_1.$$

Problèmes plus complexes

- On suppose que $X_1, \ldots, X_n$ est un échantillon aléatoire d'une distribution avec $f(x;\theta)$, où le paramètre $\theta \in \Theta \subseteq \mathbb{R}$ est inconnu. Comme précédemment, on veut tester

$$\begin{aligned} \mathcal{H}_0 &: \theta \in \Theta_0, \\ \mathcal{H}_1 &: \theta \in \Theta_1. \end{aligned}$$

- Dans la plupart des cas, l'espace des paramètres Θ contient plus de deux valeurs et on considère l'un des trois problèmes suivants :

$$\begin{aligned} \mathcal{H}_0 &: \theta \leq \theta_0, \\ \mathcal{H}_1 &: \theta > \theta_0, \end{aligned} \qquad \begin{aligned} \mathcal{H}_0 &: \theta \geq \theta_0, \\ \mathcal{H}_1 &: \theta < \theta_0, \end{aligned} \qquad \begin{aligned} \mathcal{H}_0 &: \theta = \theta_0, \\ \mathcal{H}_1 &: \theta \neq \theta_0. \end{aligned}$$

Problèmes plus complexes

- On suppose que $X_1, \ldots, X_n$ est un échantillon aléatoire d'une distribution avec $f(x;\theta)$, où le paramètre $\theta \in \Theta \subseteq \mathbb{R}$ est inconnu. Comme précédemment, on veut tester

$$\begin{aligned} \mathcal{H}_0 &: \theta \in \Theta_0, \\ \mathcal{H}_1 &: \theta \in \Theta_1. \end{aligned}$$

- Dans la plupart des cas, l'espace des paramètres Θ contient plus de deux valeurs et on considère l'un des trois problèmes suivants :

$$\begin{aligned} \mathcal{H}_0 &: \theta \leq \theta_0, & \mathcal{H}_0 &: \theta \geq \theta_0, & \mathcal{H}_0 &: \theta = \theta_0, \\ \mathcal{H}_1 &: \theta > \theta_0, & \mathcal{H}_1 &: \theta < \theta_0, & \mathcal{H}_1 &: \theta \neq \theta_0. \end{aligned}$$

- Dans ces trois cas, contre-hypothèse est composite, alors que l'hypothèse nulle est simple ou composite.

Erreur de type I revisitée

Les trois problèmes sont testés au seuil de signification spécifié α.

Erreur de type I revisitée

Les trois problèmes sont testés au seuil de signification spécifié α.

Cela signifie que

$$\Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) \leq \alpha \quad \text{pour toute valeur de } \theta \in \Theta_0,$$

ou, de façon équiavlente, que la taille du test est plus petite ou égale à α :

$$\sup_{\theta \in \Theta_0} \Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) = \sup_{\theta \in \Theta_0} \pi(\theta) \leq \alpha.$$

Erreur de type I revisitée

Les trois problèmes sont testés au seuil de signification spécifié α.

Cela signifie que

$$\Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) \leq \alpha \quad \text{pour toute valeur de } \theta \in \Theta_0,$$

ou, de façon équiavlente, que la taille du test est plus petite ou égale à α :

$$\sup_{\theta \in \Theta_0} \Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) = \sup_{\theta \in \Theta_0} \pi(\theta) \leq \alpha.$$

Cela contrôle la probabilité d'erreur de type I.

Erreur de type II revisitée

On rappelle que la fonction de puissance du test est telle que, pour tout $\theta \in \Theta_1$,

$$\pi(\theta) = \Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) = 1 - \Pr(\text{erreur de type II} \mid \theta).$$

Donc, il est désirable de maximiser $\pi(\theta)$ à chaque $\theta \in \Theta_1$.

Erreur de type II revisitée

On rappelle que la fonction de puissance du test est telle que, pour tout $\theta \in \Theta_1$,

$$\pi(\theta) = \Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) = 1 - \Pr(\text{erreur de type II} \mid \theta).$$

Donc, il est désirable de maximiser $\pi(\theta)$ à chaque $\theta \in \Theta_1$.

Pour souligner que la fonction de puissance dépend du test δ^*, on écrit

$$\pi(\theta|\delta^*) = \Pr(\text{Test } \delta^* \text{ rejette } \mathcal{H}_0 \mid \theta).$$

Erreur de type II revisitée

On rappelle que la fonction de puissance du test est telle que, pour tout $\theta \in \Theta_1$,

$$\pi(\theta) = \Pr(\text{rejeter } \mathcal{H}_0 \mid \theta) = 1 - \Pr(\text{erreur de type II} \mid \theta).$$

Donc, il est désirable de maximiser $\pi(\theta)$ à chaque $\theta \in \Theta_1$.

Pour souligner que la fonction de puissance dépend du test δ^*, on écrit

$$\pi(\theta|\delta^*) = \Pr(\text{Test } \delta^* \text{ rejette } \mathcal{H}_0 \mid \theta).$$

La propriété désirable de δ^* est donc

$$\pi(\theta|\delta^*) = \sup_{\text{test } \delta} \pi(\theta|\delta), \quad \text{pour toute valeur de } \theta \in \Theta_1.$$

Test uniformément le plus puissant

Définition

Une procédure de test δ^* est un test uniformément le plus puissant (*uniformly most powerful*, UMP) au seuil α, si

$$\sup_{\theta \in \Theta_0} \pi(\theta|\delta^*) \leq \alpha$$

et si, pour toute autre procédure de test δ telle que

$$\sup_{\theta \in \Theta_0} \pi(\theta|\delta) \leq \alpha,$$

on a

$$\pi(\theta|\delta) \leq \pi(\theta|\delta^*), \quad \text{pour toute valeur de } \theta \in \Theta_1.$$

Exemple 2

Soit $X_1, \ldots, X_n$ un échantillon de la distribution normale avec moyenne $\mu \in \mathbb{R}$ inconnue et variance σ^2 connue.

Exemple 2

Soit $X_1, \ldots, X_n$ un échantillon de la distribution normale avec moyenne $\mu \in \mathbb{R}$ inconnue et variance σ^2 connue.

On veut trouver le test optimal au seuil α pour le problème

$$\begin{aligned} \mathcal{H}_0 &: \mu = \mu_0, \\ \mathcal{H}_1 &: \mu = \mu_1, \end{aligned}$$

où $\mu_0 < \mu_1$. Ici, les deux hypothèses sont simples.

Exemple 2

Soit $X_1, \ldots, X_n$ un échantillon de la distribution normale avec moyenne $\mu \in \mathbb{R}$ inconnue et variance σ^2 connue.

On veut trouver le test optimal au seuil α pour le problème

$$\begin{aligned} \mathcal{H}_0 &: \mu = \mu_0, \\ \mathcal{H}_1 &: \mu = \mu_1, \end{aligned}$$

où $\mu_0 < \mu_1$. Ici, les deux hypothèses sont simples.

C'est une petite généralisation de l'**exemple 1**, où on avait $\sigma^2 = 1$, $\mu_0 = 0$ et $\mu_1 = 1$.

Exemple 2 (suite)

Le ratio de vraisemblance est

$$\begin{aligned}
\frac{f(\mathbf{x};\mu_1)}{f(\mathbf{x};\mu_0)} &= \exp\left\{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_1)^2\right\} \\
&= \exp\left\{\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2-2x_i\mu_0+\mu_0^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2-2x_i\mu_1+\mu_1^2)\right\} \\
&= \exp\left\{\frac{\mu_1-\mu_0}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{n(\mu_1^2-\mu_0^2)}{2\sigma^2}\right\} \\
&= \exp\left\{\frac{n(\mu_1-\mu_0)}{\sigma^2}\bar{x}_n - \frac{n(\mu_1^2-\mu_0^2)}{2\sigma^2}\right\}.
\end{aligned}$$

Exemple 2 (suite)

Selon le Lemme de Neyman–Pearson, le test optimal rejette $\mathcal{H}_0$ si, pour un $k > 0$ approprié,

$$\begin{aligned}
\frac{f(\mathbf{x};1)}{f(\mathbf{x};0)} &> \frac{1}{k} & \Leftrightarrow \\
\ln\left\{\frac{f(\mathbf{x};1)}{f(\mathbf{x};0)}\right\} &> -\ln(k) & \Leftrightarrow \\
\frac{n(\mu_1-\mu_0)}{\sigma^2}\bar{x}_n - \frac{n(\mu_1^2-\mu_0^2)}{2\sigma^2} &> -\ln(k) & \Leftrightarrow \\
\bar{x}_n &> \underbrace{\frac{\mu_1+\mu_0}{2} - \frac{\sigma^2\ln(k)}{n(\mu_1-\mu_0)}}_{=k^*}.
\end{aligned}$$

Exemple 2 (suite)

Donc, on doit trouver k ou plutôt k^* de sorte que

$$\Pr(\bar{X}_n > k^* | \mu = \mu_0) = \alpha,$$

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} > \sqrt{n}\,\frac{k^* - \mu_0}{\sigma} | \mu = \mu_0\right) = \alpha.$$

Exemple 2 (suite)

Donc, on doit trouver k ou plutôt k^* de sorte que

$$\Pr(\bar{X}_n > k^* | \mu = \mu_0) = \alpha,$$

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} > \sqrt{n}\,\frac{k^* - \mu_0}{\sigma} | \mu = \mu_0\right) = \alpha.$$

Puisque $\sqrt{n}(\bar{X}_n - \mu_0)/\sigma$ suit une normale centrée réduite sous $\mathcal{H}_0$,

$$\sqrt{n}\,\frac{k^* - \mu_0}{\sigma} = z_\alpha.$$

Exemple 2 (suite)

Donc, on doit trouver k ou plutôt k^* de sorte que

$$\Pr(\bar{X}_n > k^* | \mu = \mu_0) = \alpha,$$

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} > \sqrt{n}\,\frac{k^* - \mu_0}{\sigma} | \mu = \mu_0\right) = \alpha.$$

Puisque $\sqrt{n}(\bar{X}_n - \mu_0)/\sigma$ suit une normale centrée réduite sous $\mathcal{H}_0$,

$$\sqrt{n}\,\frac{k^* - \mu_0}{\sigma} = z_\alpha.$$

Finalement, le test optimal rejette $\mathcal{H}_0$ si

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_\alpha.$$

Exemple 2 (suite)

Incroyable ! Le test optimal ne dépend pas de μ_1 !

Exemple 2 (suite)

Incroyable ! Le test optimal ne dépend pas de μ_1 !

Pour tout $\mu_1 > \mu_0$, le test δ^* qui minimise l'erreur de type II parmi tous les tests δ au seuil α du problème

$$\mathcal{H}_0 \;:\; \mu = \mu_0,$$
$$\mathcal{H}_1 \;:\; \mu = \mu_1,$$

rejette $\mathcal{H}_0$ si

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_\alpha$$

et ne rejette pas $\mathcal{H}_0$ si

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} < z_\alpha.$$

Exemple 2 (suite)

De plus, pour tout $\mu^* < \mu_0$,

$$\begin{aligned}
\pi(\mu^*|\delta^*) &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_\alpha \middle| \mu^*\right) \\
&= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu^*}{\sigma} \geq z_\alpha + \sqrt{n}\,\frac{\mu_0 - \mu^*}{\sigma} \middle| \mu^*\right) \\
&< \alpha
\end{aligned}$$

Exemple 2 (suite)

De plus, pour tout $\mu^* < \mu_0$,

$$\begin{aligned}\pi(\mu^*|\delta^*) &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_\alpha \middle| \mu^*\right) \\ &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu^*}{\sigma} \geq z_\alpha + \sqrt{n}\,\frac{\mu_0 - \mu^*}{\sigma} \middle| \mu^*\right) \\ &< \alpha\end{aligned}$$

parce que la variable aléatoire en rouge est normale centrée réduite quand $\mu = \mu^*$ et la quantité en bleu est strictement positive lorsque $\mu^* < \mu_0$.

Exemple 2 (suite)

Donc

$$\sup_{\mu \leq \mu_0} \pi(\mu|\delta^*) \leq \alpha.$$

Exemple 2 (suite)

Donc

$$\sup_{\mu \le \mu_0} \pi(\mu|\delta^*) \le \alpha.$$

En conclusion, δ^* est le test uniformément le plus puissant au seuil α pour le problème

$$\begin{aligned} \mathcal{H}_0 &: \mu \le \mu_0, \\ \mathcal{H}_1 &: \mu > \mu_0. \end{aligned}$$

Exemple 2 (suite)

La fonction de puissance de δ^* au point $\mu_1 > \mu_0$ est

$$\begin{aligned}
\pi(\mu_1|\delta^*) &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_\alpha \middle| \mu_1\right) \\
&= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_1}{\sigma} \geq z_\alpha + \sqrt{n}\,\frac{\mu_0 - \mu_1}{\sigma} \middle| \mu_1\right) \\
&= 1 - \Phi\left(z_\alpha + \sqrt{n}\,\frac{\mu_0 - \mu_1}{\sigma}\right),
\end{aligned}$$

où Φ est la fonction de répartition de la normale centrée réduite.

Exemple 2 (suite)

La fonction de puissance de δ^* au point $\mu_1 > \mu_0$ est

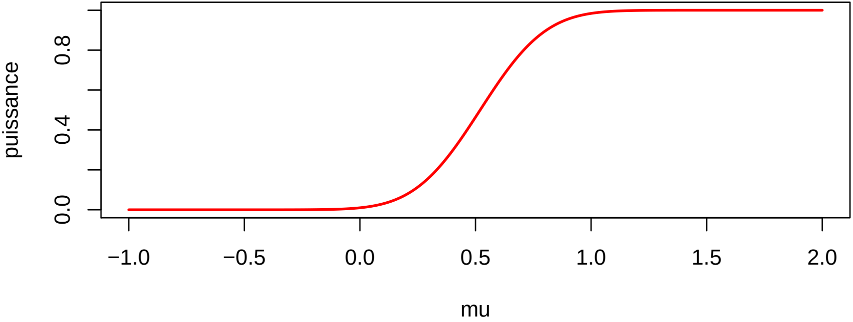
$$\begin{aligned}\pi(\mu_1|\delta^*) &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_\alpha \middle| \mu_1\right) \\ &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_1}{\sigma} \geq z_\alpha + \sqrt{n}\,\frac{\mu_0 - \mu_1}{\sigma} \middle| \mu_1\right) \\ &= 1 - \Phi\left(z_\alpha + \sqrt{n}\,\frac{\mu_0 - \mu_1}{\sigma}\right),\end{aligned}$$

où Φ est la fonction de répartition de la normale centrée réduite.

Le terme orange est négatif et tend vers $-\infty$ quand $\mu_1 \to \infty$. La fonction de puissance est donc strictement croissante et converge vers 1 quand $\mu_1 \to \infty$.

Exemple 2 (suite)

Illustration de la fonction de puissance du test uniformément le plus puissant quand $\mu_0 = 0$, $\sigma^2 = 1$, $n = 20$ et $\alpha = 0.01$.



Pourquoi cet exemple fonctionne-t-il ?

- Le rapport de vraisemblance est une fonction monotone de $\bar{X}_n$, soit

$$\frac{f(\mathbf{x};\mu_1)}{f(\mathbf{x};\mu_0)} = \exp\left\{\frac{n(\mu_1-\mu_0)}{\sigma^2}\bar{x}_n - \frac{n(\mu_1^2-\mu_0^2)}{2\sigma^2}\right\} = g_{\mu_0,\mu_1}(\bar{x}_n).$$

Pourquoi cet exemple fonctionne-t-il ?

- Le rapport de vraisemblance est une fonction monotone de $\bar{X}_n$, soit

$$\frac{f(\mathbf{x};\mu_1)}{f(\mathbf{x};\mu_0)} = \exp\left\{\frac{n(\mu_1-\mu_0)}{\sigma^2}\bar{x}_n - \frac{n(\mu_1^2-\mu_0^2)}{2\sigma^2}\right\} = g_{\mu_0,\mu_1}(\bar{x}_n).$$

- Puisque g_{μ_0,μ_1} est strictement croissante pour tout $\mu_0 < \mu_1$,

$$\frac{f(\mathbf{x};\mu_1)}{f(\mathbf{x};\mu_0)} = g_{\mu_0,\mu_1}(\bar{x}_n) > 1/k \quad \Leftrightarrow \quad \bar{x}_n > g^{-1}_{\mu_0,\mu_1}(1/k) = k^*.$$

Pourquoi cet exemple fonctionne-t-il ?

- Le rapport de vraisemblance est une fonction monotone de $\bar{X}_n$, soit

$$\frac{f(\mathbf{x};\mu_1)}{f(\mathbf{x};\mu_0)} = \exp\left\{\frac{n(\mu_1-\mu_0)}{\sigma^2}\bar{x}_n - \frac{n(\mu_1^2-\mu_0^2)}{2\sigma^2}\right\} = g_{\mu_0,\mu_1}(\bar{x}_n).$$

- Puisque g_{μ_0,μ_1} est strictement croissante pour tout $\mu_0 < \mu_1$,

$$\frac{f(\mathbf{x};\mu_1)}{f(\mathbf{x};\mu_0)} = g_{\mu_0,\mu_1}(\bar{x}_n) > 1/k \quad \Leftrightarrow \quad \bar{x}_n > g^{-1}_{\mu_0,\mu_1}(1/k) = k^*.$$

- Choisir k^* est facile puisque quand $\mu = \mu_0$, $\sqrt{n}(\bar{X}_n - \mu_0)/\sigma$ est normale centrée réduite.

Rapport de vraisemblance monotone

Définition

Une famille paramétrique $f(\mathbf{x};\theta)$, $\theta \in \Theta$ a un rapport de vraisemblance monotone en la statistique T (*monotone likelihood ratio*),

$$T = T(X_1, \ldots, X_n),$$

si pour tout $\theta_1 < \theta_2$, le rapport

$$\frac{f(\mathbf{x};\theta_2)}{f(\mathbf{x};\theta_1)} = g_{\theta_1,\theta_2}(t)$$

dépend de $(x_1, \ldots, x_n)$ seulement à travers la statistique T, $t = t(x_1, \ldots, x_n)$, et la fonction g_{θ_1,θ_2} est croissante en t.

Exemple 5

On considère un échantillon $X_1, \ldots, X_n$ de la distribution exponentielle.

Exemple 5

On considère un échantillon $X_1, \ldots, X_n$ de la distribution exponentielle.

Pour $\beta_1 < \beta_2$,

$$\frac{f(\mathbf{x}; \beta_2)}{f(\mathbf{x}; \beta_1)} = \frac{\beta_1^n}{\beta_2^n} \exp\left\{-n\bar{x}_n\left(\frac{1}{\beta_2} - \frac{1}{\beta_1}\right)\right\}$$
$$= \exp\left\{n\bar{x}_n\left(\frac{1}{\beta_1} - \frac{1}{\beta_2}\right)\right\}\left(\frac{\beta_1}{\beta_2}\right)^n.$$

Exemple 5

On considère un échantillon $X_1, \ldots, X_n$ de la distribution exponentielle.

Pour $\beta_1 < \beta_2$,

$$\begin{aligned}\frac{f(\mathbf{x};\beta_2)}{f(\mathbf{x};\beta_1)} &= \frac{\beta_1^n}{\beta_2^n}\exp\left\{-n\bar{x}_n\left(\frac{1}{\beta_2}-\frac{1}{\beta_1}\right)\right\}\\ &= \exp\left\{n\bar{x}_n\left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right)\right\}\left(\frac{\beta_1}{\beta_2}\right)^n.\end{aligned}$$

La famille a un rapport de vraisemblance monotone en

$$T = \bar{X}_n.$$

Déjà vu

Rappel: Si la statistique T est exhaustive pour θ, le **Théorème de factorisation de Fisher–Neyman** implique que

$$\frac{f(\mathbf{x};\theta_2)}{f(\mathbf{x};\theta_1)} = \frac{g(t,\theta_2)h(x_1,\ldots,x_n)}{g(t,\theta_1)h(x_1,\ldots,x_n)} = \frac{g(t,\theta_2)}{g(t,\theta_1)} = g_{\theta_1,\theta_2}(t)$$

et $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T lorsque g_{θ_1,θ_2} est croissante pour tout $\theta_1 < \theta_2$.

Déjà vu

Rappel: Si la statistique T est exhaustive pour θ, le **Théorème de factorisation de Fisher–Neyman** implique que

$$\frac{f(\mathbf{x};\theta_2)}{f(\mathbf{x};\theta_1)} = \frac{g(t,\theta_2)h(x_1,\ldots,x_n)}{g(t,\theta_1)h(x_1,\ldots,x_n)} = \frac{g(t,\theta_2)}{g(t,\theta_1)} = g_{\theta_1,\theta_2}(t)$$

et $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T lorsque g_{θ_1,θ_2} est croissante pour tout $\theta_1 < \theta_2$.

D'un autre côté, si $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T, alors T doit être exhaustive pour θ.

Test uniformément le plus puissant pour contre-hypothèse unilatérale

On considère le problème

$$\begin{aligned} \mathcal{H}_0 &: \theta \leq \theta_0, \\ \mathcal{H}_1 &: \theta > \theta_0. \end{aligned}$$

Théorème

On suppose que $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T et que c_α est une constante telle que

$$\Pr(T \geq c_\alpha | \theta = \theta_0) = \alpha.$$

La procédure de test qui rejette $\mathcal{H}_0$ si $T \geq c_\alpha$ est un test UMP du problème ci-haut au seuil de signification α.

Exemple 5 (suite)

Soit $X_1, \ldots, X_n$ un échantillon de la distribution exponentielle et on considère le problème

$$\mathcal{H}_0 \ : \ \beta \leq \beta_0,$$
$$\mathcal{H}_1 \ : \ \beta > \beta_0.$$

Exemple 5 (suite)

Soit $X_1, \ldots, X_n$ un échantillon de la distribution exponentielle et on considère le problème

$$\mathcal{H}_0 \; : \; \beta \leq \beta_0,$$
$$\mathcal{H}_1 \; : \; \beta > \beta_0.$$

Par le Théorème précédent, le test UMP au seuil α rejette $\mathcal{H}_0$ si

$$\bar{X}_n \geq c_\alpha,$$

où c est telle que

$$\Pr(\bar{X}_n \geq c_\alpha | \beta_0) = \alpha.$$

Exemple 5 (suite)

La valeur critique c_α peut être choisie en utilisant le fait que $X_1 + \cdots + X_n$ a une distribution Gamma avec paramètre de forme n et paramètre d'échelle β.

Exemple 5 (suite)

La valeur critique c_α peut être choisie en utilisant le fait que $X_1 + \cdots + X_n$ a une distribution Gamma avec paramètre de forme n et paramètre d'échelle β.

Par exemple, si $\beta_0 = 1$,

$$\Pr(\bar{X}_n \geq c_\alpha | \beta = 1) = \Pr(n\bar{X}_n \geq nc_\alpha | \beta = 1) = \Pr(W \geq nc_\alpha)$$

où $W \sim \mathcal{G}(n, 1)$. En particulier,

$$c_\alpha = w_\alpha / n$$

où w_α est le quantile d'ordre $1 - \alpha$ de la $\mathcal{G}(n, 1)$.

Exemple 5 (suite)

La valeur critique c_α peut être choisie en utilisant le fait que $X_1 + \cdots + X_n$ a une distribution Gamma avec paramètre de forme n et paramètre d'échelle β.

Par exemple, si $\beta_0 = 1$,

$$\Pr(\bar{X}_n \geq c_\alpha | \beta = 1) = \Pr(n\bar{X}_n \geq nc_\alpha | \beta = 1) = \Pr(W \geq nc_\alpha)$$

où $W \sim \mathcal{G}(n, 1)$. En particulier,

$$c_\alpha = w_\alpha / n$$

où w_α est le quantile d'ordre $1 - \alpha$ de la $\mathcal{G}(n, 1)$.

Par exemple lorsque $\alpha = 0.05$ et $n = 20$,

$$c_\alpha = \frac{27.87924}{20} = 1.39396.$$

Le seuil observé

Si $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T et que le problème est

$$\begin{aligned} \mathcal{H}_0 &: \theta \leq \theta_0, \\ \mathcal{H}_1 &: \theta > \theta_0. \end{aligned}$$

Le seuil observé

Si $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T et que le problème est

$$\begin{aligned} \mathcal{H}_0 &: \theta \leq \theta_0, \\ \mathcal{H}_1 &: \theta > \theta_0. \end{aligned}$$

Le test UMP au seuil α rejette $\mathcal{H}_0$ si $T \geq c_\alpha$ pour une valeur critique c_α choisie.

Le seuil observé

Si $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T et que le problème est

$$\begin{aligned}\mathcal{H}_0 &: \theta \leq \theta_0,\\ \mathcal{H}_1 &: \theta > \theta_0.\end{aligned}$$

Le test UMP au seuil α rejette $\mathcal{H}_0$ si $T \geq c_\alpha$ pour une valeur critique c_α choisie.

Le seuil observé du test est donné par

$$\Pr(T \geq t | \theta = \theta_0),$$

où t est la valeur observée de la statistique du test T.

Exemple 5 (suite)

Supposons que dans l'**exemple 5** le problème est

$$\begin{aligned} \mathcal{H}_0 &: \beta \leq 1, \\ \mathcal{H}_1 &: \beta > 1, \end{aligned}$$

et que pour un échantillon de taille $n = 20$, $\bar{x}_n = 1.28$.

Exemple 5 (suite)

Supposons que dans l'**exemple 5** le problème est

$$\mathcal{H}_0 \ : \ \beta \leq 1,$$
$$\mathcal{H}_1 \ : \ \beta > 1,$$

et que pour un échantillon de taille $n = 20$, $\bar{x}_n = 1.28$.

Le seuil observé est donc

$$\Pr(\bar{X}_n \geq 1.28|\beta = 1)$$

et est égal à

```
1-pgamma(1.28*20,shape=20,scale=1)

## [1] 0.1104059
```

Remarque

Le Théorème sur les tests UMP peut être étendu aux problèmes de la forme

$$\mathcal{H}_0 \ : \ \theta \geq \theta_0,$$
$$\mathcal{H}_1 \ : \ \theta < \theta_0.$$

Remarque

Le Théorème sur les tests UMP peut être étendu aux problèmes de la forme

$$\begin{aligned} \mathcal{H}_0 &: \theta \geq \theta_0, \\ \mathcal{H}_1 &: \theta < \theta_0. \end{aligned}$$

Si $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T et c_α est une constante telle que

$$\Pr(T \leq c_{1-\alpha} | \theta = \theta_0) = \alpha,$$

la procédure de test qui rejette $\mathcal{H}_0$ si $T \leq c_{1-\alpha}$ est un test UMP de ce problème au seuil de signification α.

Remarque

Le Théorème sur les tests UMP peut être étendu aux problèmes de la forme

$$\begin{aligned}\mathcal{H}_0 &: \theta \geq \theta_0, \\ \mathcal{H}_1 &: \theta < \theta_0.\end{aligned}$$

Si $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T et c_α est une constante telle que

$$\Pr(T \leq c_{1-\alpha}|\theta = \theta_0) = \alpha,$$

la procédure de test qui rejette $\mathcal{H}_0$ si $T \leq c_{1-\alpha}$ est un test UMP de ce problème au seuil de signification α. Le seuil observé de ce test est

$$\Pr(T \leq t|\theta = \theta_0),$$

où t est la valeur observée de la statistique du test T.

Test UMP pour contre-hypothèse bilatérale

Observation

Il n'existe pas de test UMP pour un problème avec une contre-hypothèse bilatérale.

Test UMP pour contre-hypothèse bilatérale

Observation

Il n'existe pas de test UMP pour un problème avec une contre-hypothèse bilatérale.

En effet, supposons que $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T. Alors:

Test UMP pour contre-hypothèse bilatérale

Observation

Il n'existe pas de test UMP pour un problème avec une contre-hypothèse bilatérale.

En effet, supposons que $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T. Alors:

- Pour $\theta_1 > \theta_0$, la puissance est maximisée par le test δ_1 qui rejette $\mathcal{H}_0$ si $T \geq c_\alpha$. Par contre, pour tout $\theta_2 < \theta_0$,

$$\pi(\theta_2|\delta_1) \leq \alpha.$$

Test UMP pour contre-hypothèse bilatérale

Observation

Il n'existe pas de test UMP pour un problème avec une contre-hypothèse bilatérale.

En effet, supposons que $f(\mathbf{x};\theta)$ a un rapport de vraisemblance monotone en T. Alors:

- Pour $\theta_1 > \theta_0$, la puissance est maximisée par le test δ_1 qui rejette $\mathcal{H}_0$ si $T \geq c_\alpha$. Par contre, pour tout $\theta_2 < \theta_0$,

$$\pi(\theta_2|\delta_1) \leq \alpha.$$

- Pour $\theta_1 < \theta_0$, la puissance est maximisée par le test δ_2 qui rejette $\mathcal{H}_0$ si $T \leq c_{1-\alpha}$. Par contre, pour tout $\theta_2 > \theta_0$,

$$\pi(\theta_2|\delta_2) \leq \alpha.$$

Un compromis raisonnable

Il semble raisonnable d'utiliser un test δ qui rejette $\mathcal{H}_0$ si $T \leq c_{1-\alpha_1}$ ou $T \geq c_{\alpha_2}$ pour des valeurs critiques choisies α_1 et α_2.

Un compromis raisonnable

Il semble raisonnable d'utiliser un test δ qui rejette $\mathcal{H}_0$ si $T \leq c_{1-\alpha_1}$ ou $T \geq c_{\alpha_2}$ pour des valeurs critiques choisies α_1 et α_2.

Pour obtenir un test au seuil α, les valeurs critiques $c_{1-\alpha_1}$ et c_{α_2} sont telles que

$$\Pr(T \leq c_{1-\alpha_1} | \theta = \theta_0) = \alpha_1, \quad \Pr(T \geq c_{\alpha_2} | \theta = \theta_0) = \alpha_2$$

et

$$\alpha_1 + \alpha_2 = \alpha.$$

Un compromis raisonnable

Il semble raisonnable d'utiliser un test δ qui rejette $\mathcal{H}_0$ si $T \leq c_{1-\alpha_1}$ ou $T \geq c_{\alpha_2}$ pour des valeurs critiques choisies α_1 et α_2.

Pour obtenir un test au seuil α, les valeurs critiques $c_{1-\alpha_1}$ et c_{α_2} sont telles que

$$\Pr(T \leq c_{1-\alpha_1} | \theta = \theta_0) = \alpha_1, \quad \Pr(T \geq c_{\alpha_2} | \theta = \theta_0) = \alpha_2$$

et

$$\alpha_1 + \alpha_2 = \alpha.$$

Le plus souvent, on choisit, par convention

$$\alpha_1 = \alpha_2 = \alpha/2.$$

Exemple 2 (suite)

On considère un échantillon aléatoire $X_1, \ldots, X_n$ d'une distribution normale avec moyenne μ inconnue et variance σ^2 connue. On veut tester

$$\mathcal{H}_0 : \mu = \mu_0,$$
$$\mathcal{H}_1 : \mu \neq \mu_0.$$

Exemple 2 (suite)

On considère un échantillon aléatoire $X_1, \ldots, X_n$ d'une distribution normale avec moyenne μ inconnue et variance σ^2 connue. On veut tester

$$\begin{aligned} \mathcal{H}_0 &: \mu = \mu_0, \\ \mathcal{H}_1 &: \mu \neq \mu_0. \end{aligned}$$

La statistique du test est encore

$$T = \bar{X}_n$$

puisque la famille normale a un rapport de vraisemblance monotone en T.

Exemple 2 (suite)

Pour développer un test acceptable, on note que sous l'hypothèse nulle,

$$\sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma}$$

suit une normale centrée réduite.

Exemple 2 (suite)

Pour développer un test acceptable, on note que sous l'hypothèse nulle,

$$\sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma}$$

suit une normale centrée réduite. Si on pose

$$c_{1-\alpha_1} = \mu_0 - z_{\alpha_1}\frac{\sigma}{\sqrt{n}}, \qquad c_{\alpha_2} = \mu_0 + z_{\alpha_2}\frac{\sigma}{\sqrt{n}}$$

pour $\alpha_1, \alpha_2 \geq 0$ tels que $\alpha_1 + \alpha_2 = \alpha$,

Exemple 2 (suite)

Pour développer un test acceptable, on note que sous l'hypothèse nulle,

$$\sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma}$$

suit une normale centrée réduite. Si on pose

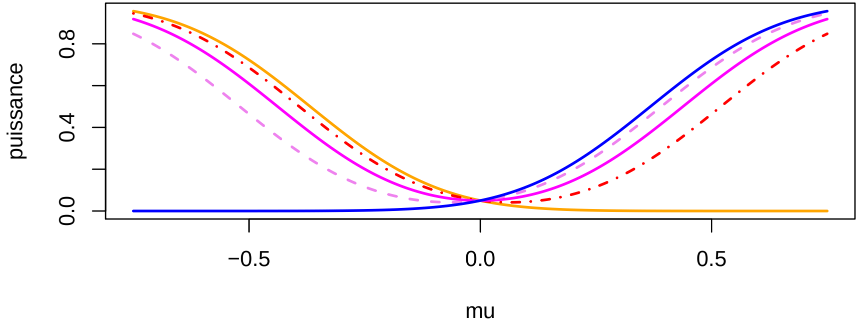
$$c_{1-\alpha_1} = \mu_0 - z_{\alpha_1}\frac{\sigma}{\sqrt{n}}, \qquad c_{\alpha_2} = \mu_0 + z_{\alpha_2}\frac{\sigma}{\sqrt{n}}$$

pour $\alpha_1, \alpha_2 \geq 0$ tels que $\alpha_1 + \alpha_2 = \alpha$, alors

$$\begin{aligned}
&\Pr(\bar{X}_n \leq c_{1-\alpha_1} | \mu = \mu_0) + \Pr(\bar{X}_n \geq c_{\alpha_2} | \mu = \mu_0) \\
&= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \leq -z_{\alpha_1}\right) + \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\sigma} \geq z_{\alpha_2}\right) \\
&= \alpha_1 + \alpha_2 = \alpha.
\end{aligned}$$

Exemple 2 (la fin!!!!)

Le graphique ci-dessous montre la fonction de puissance du test quand $\mu_0 = 0$, $\sigma^2 = 1$, $\alpha = 0.05$ et $\alpha_1 = 0.05$, $\alpha_2 = 0$, $\alpha_1 = 0.04$, $\alpha_2 = 0.01$, $\alpha_1 = \alpha_2 = 0.025$, $\alpha_1 = 0.01$, $\alpha_2 = 0.04$, et $\alpha_1 = 0$, $\alpha_2 = 0.05$.



Tests avec petits échantillons de loi normale

Dans cette section, on suppose qu'on a l'échantillon aléatoire

$$X_1, \ldots, X_n$$

tiré de la distribution normale avec moyenne μ et variance σ^2.

Tests avec petits échantillons de loi normale

Dans cette section, on suppose qu'on a l'échantillon aléatoire

$$X_1, \ldots, X_n$$

tiré de la distribution normale avec moyenne μ et variance σ^2.

Les problèmes suivants peuvent être intéressants :

- test d'hypothèse sur la moyenne μ
- test d'hypothèse sur le variance σ^2
- test d'hypothèses sur la différence de moyennes basé sur deux échantillons indépendants.

Exemple

Le thon rouge est l'espèce de thon la plus grande et la plus menacée ; cette espèce a été trop pêchée et on pense que le poids moyen a diminué. On veut tester si le poids moyen est inférieur à 550 livres.

Exemple

Le thon rouge est l'espèce de thon la plus grande et la plus menacée ; cette espèce a été trop pêchée et on pense que le poids moyen a diminué. On veut tester si le poids moyen est inférieur à 550 livres.

Exemple (suite)

Soit l'échantillon aléatoire des poids de $n = 12$ thons rouges indépendants obtenus de bateaux commerciaux. Les poids sont normalement distribués et on a

$$\bar{x}_n = 535.7 \quad \text{et} \quad s_n = 37.8.$$

Peut-on conclure que le poids moyen est sous 550 livres ?

Si seulement on connaissait la variance !

On veut faire un test unilatéral sur la moyenne

$$\begin{aligned} \mathcal{H}_0 &: \mu \leq (\geq) \mu_0, \\ \mathcal{H}_1 &: \mu > (<) \mu_0 \end{aligned}$$

Si seulement on connaissait la variance !

On veut faire un test unilatéral sur la moyenne

$$\mathcal{H}_0 : \mu \leq (\geq) \mu_0,$$
$$\mathcal{H}_1 : \mu > (<) \mu_0$$

Si la variance σ^2 était connue, alors le test uniformément le plus puissant au seuil α rejette $\mathcal{H}_0$ si

$$\sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma} \geq (\leq -) z_\alpha,$$

où z_α est le quantile $1 - \alpha$ de la distribution normale centrée réduite.

Pas de test uniformément le plus puissant

- Quand σ^2 est inconnu, on peut montrer qu'il n'y a pas de test uniformément le plus puissant.

Pas de test uniformément le plus puissant

- Quand σ^2 est inconnu, on peut montrer qu'il n'y a pas de test uniformément le plus puissant.
- Idée: Commençons avec la statistique du test optimal quand σ^2 est connu et estimons σ^2 par la variance échantillonnale S_n^2.

Pas de test uniformément le plus puissant

- Quand σ^2 est inconnu, on peut montrer qu'il n'y a pas de test uniformément le plus puissant.

- Idée: Commençons avec la statistique du test optimal quand σ^2 est connu et estimons σ^2 par la variance échantillonnale S_n^2.

- Pour trouver la région critique, on utilise le fait que quand $\mu = \mu_0$,

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n}$$

suit une loi de Student t avec $n-1$ degrés de liberté.

Test t unilatéral

Test t unilatéral

Le test t au seuil α du problème avec une contre-hypothèse unilatérale

$$\mathcal{H}_0 \ : \ \mu \leq (\geq) \mu_0,$$
$$\mathcal{H}_1 \ : \ \mu > (<) \mu_0$$

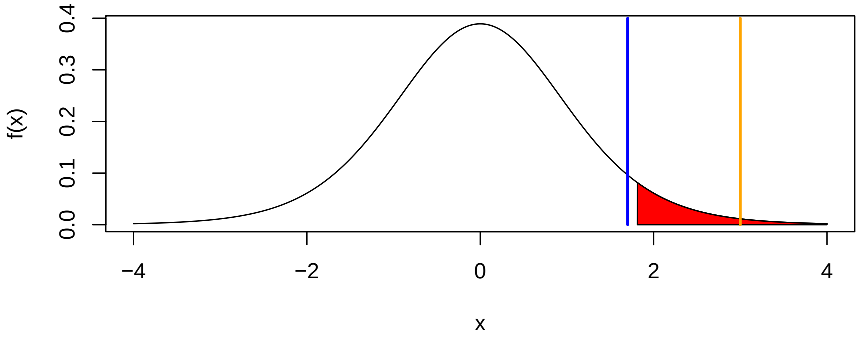
rejette l'hypothèse nulle $\mathcal{H}_0$ si

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \geq (\leq -) t_{n-1,\alpha},$$

où $t_{n-1,\alpha}$ est le quantile $1-\alpha$ de la loi t avec $n-1$ dl.

Illustration

Région critique (en rouge) du Test t unilatéral avec $\mathcal{H}_0 : \mu \leq \mu_0$ quand $n = 11$ et $\alpha = 0.05$. La ligne bleue est une valeur observée de la statistique du test pour laquelle $\mathcal{H}_0$ n'est pas rejetée, la ligne orange est une valeur observée de la statistique du test pour laquelle $\mathcal{H}_0$ est rejetée.



Taille du test t unilatéral

Pour simplifier, on se concentre sur $\mathcal{H}_0 : \mu \leq \mu_0$ versus $\mathcal{H}_1 : \mu > \mu_0$; le problème avec l'inégalité inverse peut être traité de la même façon.

Taille du test t unilatéral

Pour simplifier, on se concentre sur $\mathcal{H}_0 : \mu \leq \mu_0$ versus $\mathcal{H}_1 : \mu > \mu_0$; le problème avec l'inégalité inverse peut être traité de la même façon.

La taille du test est en effet α, car

$$\Pr\left(\sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha} \middle| \mu = \mu_0\right) = \alpha$$

Taille du test t unilatéral

Pour simplifier, on se concentre sur $\mathcal{H}_0 : \mu \leq \mu_0$ versus $\mathcal{H}_1 : \mu > \mu_0$; le problème avec l'inégalité inverse peut être traité de la même façon.

La taille du test est en effet α, car

$$\Pr\left(\sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha} \middle| \mu = \mu_0\right) = \alpha$$

et pour tout $\mu^* < \mu_0$,

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha} \middle| \mu = \mu^*\right)$$
$$= \Pr\left(\underbrace{\sqrt{n}\,\frac{\bar{X}_n - \mu^*}{S_n}}_{\sim t_{n-1}} \geq t_{n-1,\alpha} + \underbrace{\sqrt{n}\,\frac{\mu_0 - \mu^*}{S_n}}_{>0} \middle| \mu = \mu^*\right) < \alpha.$$

Puissance du test t unilatéral

Pour calculer la puissance du test t, on choisit d'abord une valeur arbitraire $\mu^* > \mu_0$. Alors,

$$\pi(\mu^*) = \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha} \middle| \mu = \mu^*\right)$$

$$= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu^*}{S_n} \geq t_{n-1,\alpha} + \underbrace{\sqrt{n}\,\frac{\mu_0 - \mu^*}{S_n}}_{<0} \middle| \mu = \mu^*\right) > \alpha.$$

Puissance du test t unilatéral

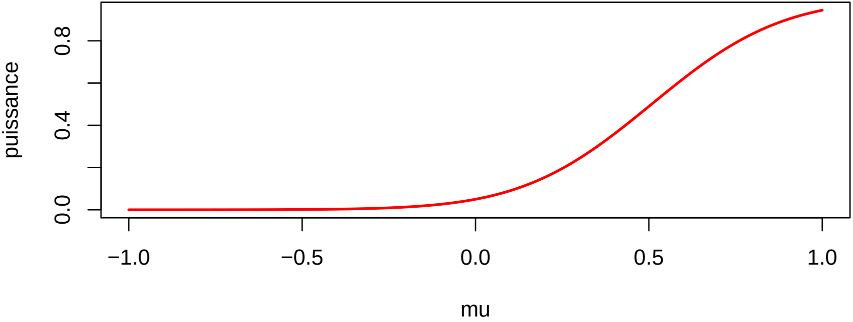
Pour calculer la puissance du test t, on choisit d'abord une valeur arbitraire $\mu^* > \mu_0$. Alors,

$$\begin{aligned}\pi(\mu^*) &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha}\middle|\mu = \mu^*\right)\\ &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu^*}{S_n} \geq t_{n-1,\alpha} + \underbrace{\sqrt{n}\,\frac{\mu_0 - \mu^*}{S_n}}_{<0}\middle|\mu = \mu^*\right) > \alpha.\end{aligned}$$

De plus, quand $\mu^* \to \infty$, le terme du côté droit dans la probabilité tend vers $-\infty$, et la puissance du test tend vers 1.

Puissance du test t unilatéral

Le graphique suivant montre la puissance du test t unilatéral pour $\mathcal{H}_0 : \mu \leq 0$ versus $\mathcal{H}_1 : \mu > 0$ quand $n = 20$, $s_n = 1.28$, et $\alpha = 0.05$.



Exemple

Dans l'exemple du thon rouge, on teste

$$\mathcal{H}_0 \ : \ \mu \geq 550, \qquad \mathcal{H}_1 \ : \ \mu < 550.$$

Exemple

Dans l'exemple du thon rouge, on teste

$$\mathcal{H}_0 \ : \ \mu \geq 550, \qquad \mathcal{H}_1 \ : \ \mu < 550.$$

La statistique t de Student est

$$t_n = \sqrt{12}\,\frac{\bar{x}_n - 550}{s_n} = \sqrt{12}\,\frac{535.7 - 550}{37.8} = -1.31049.$$

Exemple

Dans l'exemple du thon rouge, on teste

$$\mathcal{H}_0 \ : \ \mu \geq 550, \qquad \mathcal{H}_1 \ : \ \mu < 550.$$

La statistique t de Student est

$$t_n = \sqrt{12}\,\frac{\bar{x}_n - 550}{s_n} = \sqrt{12}\,\frac{535.7 - 550}{37.8} = -1.31049.$$

Si on veut tester au seuil $\alpha = 0.05$, on trouve que

$$t_{n-1,\alpha} = t_{11,0.05} = 1.796.$$

Comme

$$t_n = -1.31049 > -t_{n-1,\alpha} = -1.796,$$

Exemple

Dans l'exemple du thon rouge, on teste

$$\mathcal{H}_0 \ : \ \mu \geq 550, \qquad \mathcal{H}_1 \ : \ \mu < 550.$$

La statistique t de Student est

$$t_n = \sqrt{12}\,\frac{\bar{x}_n - 550}{s_n} = \sqrt{12}\,\frac{535.7 - 550}{37.8} = -1.31049.$$

Si on veut tester au seuil $\alpha = 0.05$, on trouve que

$$t_{n-1,\alpha} = t_{11,0.05} = 1.796.$$

Comme

$$t_n = -1.31049 > -t_{n-1,\alpha} = -1.796,$$

on ne peut pas rejeter l'hypothèse nulle au seuil 5 %. On ne peut pas conclure que le poids moyen est sous 550 livres.

Seuil observé du test t unilatéral

On considère le problème

$$\begin{aligned} \mathcal{H}_0 &: \mu \leq (\geq) \mu_0, \\ \mathcal{H}_1 &: \mu > (<) \mu_0 \end{aligned}$$

et on suppose que la valeur observée de la statistique t est

$$t_n = \sqrt{n}\,\frac{\bar{x}_n - \mu_0}{s_n}.$$

Seuil observé du test t unilatéral

On considère le problème

$$\mathcal{H}_0 \ : \ \mu \leq (\geq) \mu_0,$$
$$\mathcal{H}_1 \ : \ \mu > (<) \mu_0$$

et on suppose que la valeur observée de la statistique t est

$$t_n = \sqrt{n}\,\frac{\bar{x}_n - \mu_0}{s_n}.$$

Seuil observé du test t unilatéral

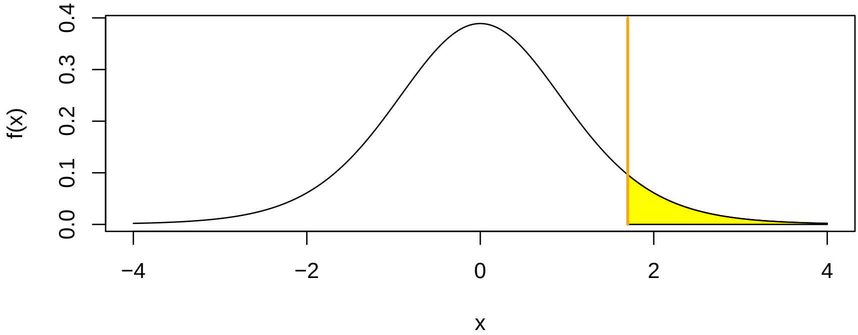
Le seuil observé du test t est

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \geq (\leq) t_n \middle| \mu = \mu_0\right) = \Pr(T_{n-1} \geq (\leq) t_n)$$

où T_{n-1} est une variable aléatoire Student t avec $n-1$ dl.

Illustration du seuil observé

Seuil observé (en jaune) du test t unilatéral pour $\mathcal{H}_0 : \mu \leq \mu_0$ quand $n = 11$. La ligne verticale orange est la valeur observée de la statistique du test.



Exemple

Le seuil observé dans l'exemple du thon rouge est

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - 550}{S_n} \leq -1.31049 \middle| \mu = 550\right).$$

Exemple

Le seuil observé dans l'exemple du thon rouge est

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - 550}{S_n} \leq -1.31049 \middle| \mu = 550\right).$$

Sachant que $\mu = 550$, la quantité en bleu suit une Student t avec 11 dl.

Le seuil observé est donc

```
(t.obs <- sqrt(12)*((535.7 - 550)/37.8))

## [1] -1.310493

pt(t.obs, df=11)

## [1] 0.1083658
```

Exemple

Le seuil observé dans l'exemple du thon rouge est

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n - 550}{S_n} \leq -1.31049 \middle| \mu = 550\right).$$

Sachant que $\mu = 550$, la quantité en bleu suit une Student t avec 11 dl.

Le seuil observé est donc

```
(t.obs <- sqrt(12)*((535.7 - 550)/37.8))

## [1] -1.310493

pt(t.obs, df=11)

## [1] 0.1083658
```

On ne peut pas rejeter $\mathcal{H}_0$ à n'importe quel seuil $\alpha \leq 0.10837$.

Exemple (suite)

De façon alternative, on pourrait construire intervalle de confiance unilatéral à droite au niveau 95% pour μ:

$$\left(-\infty, \bar{X}_n + t_{n-1,\alpha}\frac{S_n}{\sqrt{n}}\right],$$

c'est-à-dire

$$\left(-\infty, 535.7 + 1.796\frac{37.8}{\sqrt{12}}\right] = (-\infty, 555.298].$$

Exemple (suite)

De façon alternative, on pourrait construire intervalle de confiance unilatéral à droite au niveau 95% pour μ:

$$\left(-\infty, \bar{X}_n + t_{n-1,\alpha}\frac{S_n}{\sqrt{n}}\right],$$

c'est-à-dire

$$\left(-\infty, 535.7 + 1.796\frac{37.8}{\sqrt{12}}\right] = (-\infty, 555.298].$$

Comme la valeur 550 est contenu dans l'intervalle, il n'y a pas assez d'information dans les données pour conclure, au seuil de 5%, que le poids moyen du thon rouge est inférieur à 550 livres.

Test t unilatéral et intervalle de confiance unilatéral

On note que

$$\mu_0 \notin \left(-\infty, \bar{X}_n + t_{n-1,\alpha}\frac{S_n}{\sqrt{n}}\right]$$

si et seulement si

$$\mu_0 > \bar{X}_n + t_{n-1,\alpha}\frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} < -t_{n-1,\alpha}.$$

Test t unilatéral et intervalle de confiance unilatéral

On note que

$$\mu_0 \notin \left(-\infty, \bar{X}_n + t_{n-1,\alpha} \frac{S_n}{\sqrt{n}}\right]$$

si et seulement si

$$\mu_0 > \bar{X}_n + t_{n-1,\alpha} \frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} < -t_{n-1,\alpha}.$$

En bref, le test t du problème unilatéral

$$\mathcal{H}_0 \,:\, \mu \geq \mu_0, \quad \mathcal{H}_1 \,:\, \mu < \mu_0$$

rejette l'hypothèse nulle si et seulement si μ_0 n'est pas dans l'intervalle de confiance unilatéral à droite pour μ avec niveau de confiance $100 \times (1-\alpha)\%$.

Test t unilatéral et intervalle de confiance unilatéral (suite)

De façon similaire,

$$\mu_0 \notin \left[\bar{X}_n - t_{n-1,\alpha}\frac{S_n}{\sqrt{n}}, \infty\right)$$

si et seulement si

$$\mu_0 < \bar{X}_n - t_{n-1,\alpha}\frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} > t_{n-1,\alpha}.$$

Test t unilatéral et intervalle de confiance unilatéral (suite)

De façon similaire,

$$\mu_0 \notin \left[\bar{X}_n - t_{n-1,\alpha}\frac{S_n}{\sqrt{n}}, \infty\right)$$

si et seulement si

$$\mu_0 < \bar{X}_n - t_{n-1,\alpha}\frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} > t_{n-1,\alpha}.$$

En bref, le test t du problème unilatéral

$$\mathcal{H}_0 \;:\; \mu \leq \mu_0, \quad \mathcal{H}_1 \;:\; \mu > \mu_0$$

rejette l'hypothèse nulle si et seulement si μ_0 n'est pas dans l'intervalle de confiance unilatéral à gauche pour μ avec niveau de confiance $100 \times (1-\alpha)\%$.

Exemple

Malgré un système de recyclage sophistiqué, un parc aquatique informe le service des eaux de la ville du besoin de 1 million de litres d'eau par jour.

Exemple

Malgré un système de recyclage sophistiqué, un parc aquatique informe le service des eaux de la ville du besoin de 1 million de litres d'eau par jour.

Le service des eaux de la ville a sélectionné un échantillon aléatoire de $n = 21$ jours; la moyenne échantillonnale et l'écart-type échantillonnal de la consommation d'eau (en millier de litres) sont

$$\bar{x}_n = 927.43, \quad \text{et} \quad s_n = 154.45.$$

Exemple

Malgré un système de recyclage sophistiqué, un parc aquatique informe le service des eaux de la ville du besoin de 1 million de litres d'eau par jour.

Le service des eaux de la ville a sélectionné un échantillon aléatoire de $n = 21$ jours; la moyenne échantillonnale et l'écart-type échantillonnal de la consommation d'eau (en millier de litres) sont

$$\bar{x}_n = 927.43, \quad \text{et} \quad s_n = 154.45.$$

En supposant que la consommation est normalement distribuée, peut-on conclure que la consommation d'eau est différente de 1 million de litres par jour ?

Tester une contre-hypothèse bilatérale

On considère le problème

$$\begin{aligned}\mathcal{H}_0 &: \mu = \mu_0, \\ \mathcal{H}_1 &: \mu \neq \mu_0.\end{aligned}$$

Tester une contre-hypothèse bilatérale

On considère le problème

$$\mathcal{H}_0 \ : \ \mu = \mu_0,$$
$$\mathcal{H}_1 \ : \ \mu \neq \mu_0.$$

Pour construire un test t bilatéral pour ce problème, on considère encore la statistique du test

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n}.$$

Sous l'hypothèse nulle, sa valeur devrait être proche de 0.

Test t bilatéral

Test t bilatéral

Le test t au seuil α du problème avec une contre-hypothèse bilatérale

$$\mathcal{H}_0 \ : \ \mu = \mu_0,$$
$$\mathcal{H}_1 \ : \ \mu \neq \mu_0$$

Test t bilatéral

Test t bilatéral

Le test t au seuil α du problème avec une contre-hypothèse bilatérale

$$\begin{aligned} \mathcal{H}_0 &: \mu = \mu_0, \\ \mathcal{H}_1 &: \mu \neq \mu_0 \end{aligned}$$

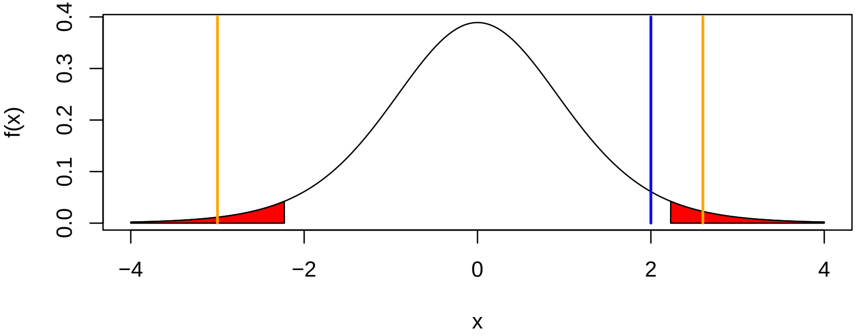
rejette l'hypothèse nulle $\mathcal{H}_0$ si

$$\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \leq -t_{n-1,\alpha/2} \quad \text{ou} \quad \sqrt{n}\,\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha/2};$$

$t_{n-1,\alpha/2}$ est le quantile d'ordre $1 - \alpha/2$ de la $t_{(n-1)}$.

Illustration de la région critique

Région critique (en rouge) du test t bilatéral quand $n = 11$ et $\alpha = 0.05$. La ligne bleue est une valeur observée de la statistique du test pour laquelle on ne rejette pas $\mathcal{H}_0$, les lignes oranges sont des valeurs observées de la statistique du test pour lesquelles on rejette $\mathcal{H}_0$.



Exemple

Dans l'exemple du parc aquatique, on veut tester au seuil $\alpha = 0.05$.

Exemple

Dans l'exemple du parc aquatique, on veut tester au seuil $\alpha = 0.05$.

Le 0.975e quantile de la distribution $t_{(n-1)} = t_{(20)}$ est

$$t_{20,0.025} = 2.086.$$

Exemple

Dans l'exemple du parc aquatique, on veut tester au seuil $\alpha = 0.05$.

Le 0.975e quantile de la distribution $t_{(n-1)} = t_{(20)}$ est

$$t_{20,0.025} = 2.086.$$

La valeur observée de la statistique t est

$$t_n = \sqrt{n}\,\frac{\bar{x}_n - 1000}{s_n} = \sqrt{21}\,\frac{927.43 - 1000}{154.45} = -2.15317.$$

Comme

$$t_n = -2.15317 < -2.086,$$

l'hypothèse que le parc a besoin d'un million de litres d'eau par jour est rejetée au seuil de 5%.

Taille du test t bilatéral

Clairement, la taille du test t bilatéral est bien α, puisque

$$\Pr\left(\sqrt{n}\,\frac{\bar{X}_n-\mu_0}{S_n}\leq -t_{n-1,\alpha/2}\middle|\mu=\mu_0\right) + \Pr\left(\sqrt{n}\,\frac{\bar{X}_n-\mu_0}{S_n}\geq t_{n-1,\alpha/2}\middle|\mu=\mu_0\right)=\alpha,$$

sachant que la quantité en bleu suit une distribution $t_{(n-1)}$ quand $\mu=\mu_0$.

Puissance du test t bilatéral

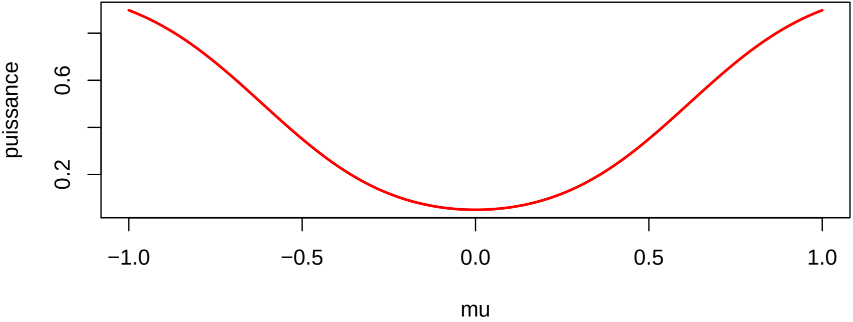
La puissance du test t bilatéral lorsque $\mu_1 \neq \mu_0$ est

$$\begin{aligned}
\pi(\mu_1) &= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n-\mu_0}{S_n} \leq -t_{n-1,\alpha/2} \middle| \mu=\mu_1\right) \\
&\quad + \Pr\left(\sqrt{n}\,\frac{\bar{X}_n-\mu_0}{S_n} \geq t_{n-1,\alpha/2} \middle| \mu=\mu_1\right) \\
&= \Pr\left(\sqrt{n}\,\frac{\bar{X}_n-\mu_1}{S_n} \leq -t_{n-1,\alpha/2} + \sqrt{n}\,\frac{\mu_0-\mu_1}{S_n} \middle| \mu=\mu_1\right) \\
&\quad + \Pr\left(\sqrt{n}\,\frac{\bar{X}_n-\mu_1}{S_n} \geq t_{n-1,\alpha/2} + \sqrt{n}\,\frac{\mu_0-\mu_1}{S_n} \middle| \mu=\mu_1\right) \\
&= 1 - \Pr\left(-t_{n-1,\alpha/2} + \sqrt{n}\,\frac{\mu_0-\mu_1}{S_n} \leq T_{n-1} \leq t_{n-1,\alpha/2} + \sqrt{n}\,\frac{\mu_0-\mu_1}{S_n}\right)
\end{aligned}$$

où T_{n-1} suit une distribution $t_{(n-1)}$. On note que quand $\mu=\mu_1$, la quantité en bleu suit une $t_{(n-1)}$.

Illustration de la puissance du test t bilatéral

Le graphique montre la puissance du test t bilatéral pour un échantillon de taille $n = 20$ avec $s_n = 1.28$ quand $\alpha = 0.05$ et $\mu_0 = 0$.



Seuil observé du test t bilatéral

Soit

$$t_n = \sqrt{n}\,\frac{\bar{x}_n - \mu_0}{s_n}$$

la valeur observée de la statistique du test t.

Seuil observé du test t bilatéral

Soit

$$t_n = \sqrt{n}\,\frac{\bar{x}_n - \mu_0}{s_n}$$

la valeur observée de la statistique du test t.

Seuil observé du test t bilatéral

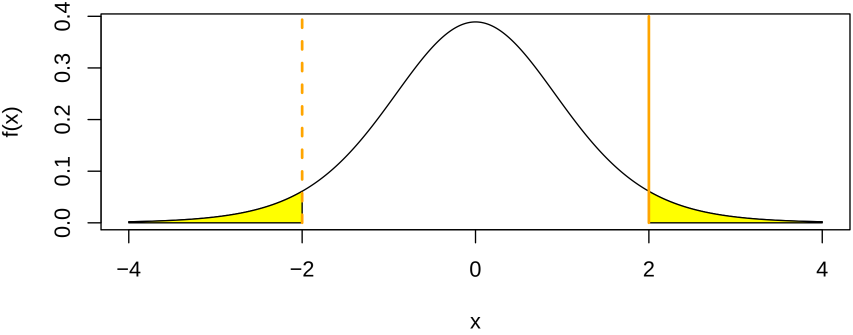
Le seuil observé du test t bilatéral est donné par

$$\begin{aligned}\Pr(|T_{n-1}| \geq |t_n|) = \Pr(T_{n-1} \geq |t_n|) + \Pr(T_{n-1} \leq -|t_n|) \\ = 2\Pr(T_{n-1} \geq |t_n|),\end{aligned}$$

où T_{n-1} est une variable aléatoire Student $t_{(n-1)}$.

Illustration du seuil observé

Seuil observé (en jaune) du test t bilatéral quand $n = 11$. La ligne orange est la valeur observée de la statistique du test, la ligne orange pointillée est la valeur observée de la statistique du test multipliée par -1.



Exemple

Dans l'exemple du parc aquatique,

$$t_n = -2.15317.$$

Le seuil observé du test bilatéral est

$$2\,\Pr(T_{n-1} \geq 2.15317).$$

Avec R, on trouve

```r
2*(1- pt(-sqrt(21)*(927.43-1000)/154.45, df=20))

## [1] 0.04368865
```

Bref, $\mathcal{H}_0$ est rejetée au seuil $\alpha = 0.05$, mais pas au seuil $\alpha = 0.01$.

Exemple (suite)

On pourrait aussi calculer l'intervalle de confiance bilatéral pour μ :

$$\left[\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right].$$

Exemple (suite)

On pourrait aussi calculer l'intervalle de confiance bilatéral pour μ :

$$\left[\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right].$$

L'intervalle de confiance à 95% est

$$[855.389, 999.471] \ ;$$

cet intervalle ne contient pas la valeur $\mu_0 = 1000$.

Exemple (suite)

On pourrait aussi calculer l'intervalle de confiance bilatéral pour μ :

$$\left[\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right].$$

L'intervalle de confiance à 95% est

$$[855.389, 999.471] \ ;$$

cet intervalle ne contient pas la valeur $\mu_0 = 1000$. Par contre, l'intervalle de confiance à 99%

$$[829.163, 1025.697];$$

inclut cette valeur, et on ne peut pas conclure que la consommation moyenne d'eau est différente de 1 million de litres à ce niveau de confiance.

L'intervalle de confiance bilatéral et le test t bilatéral

On observe que

$$\mu_0 \notin \left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}\right)$$

L'intervalle de confiance bilatéral et le test t bilatéral

On observe que

$$\mu_0 \notin \left(\bar{X}_n - t_{n-1,\alpha/2} \frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{n-1,\alpha/2} \frac{S_n}{\sqrt{n}} \right)$$

si et seulement si soit

$$\mu_0 \leq \bar{X}_n - t_{n-1,\alpha/2} \frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n} \frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha/2}.$$

ou

$$\mu_0 \geq \bar{X}_n + t_{n-1,\alpha/2} \frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n} \frac{\bar{X}_n - \mu_0}{S_n} \leq -t_{n-1,\alpha/2}.$$

L'intervalle de confiance bilatéral et le test t bilatéral

On observe que

$$\mu_0 \notin \left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}\right)$$

si et seulement si soit

$$\mu_0 \leq \bar{X}_n - t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} \geq t_{n-1,\alpha/2}.$$

ou

$$\mu_0 \geq \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}} \quad \Leftrightarrow \quad \sqrt{n}\frac{\bar{X}_n - \mu_0}{S_n} \leq -t_{n-1,\alpha/2}.$$

En bref, le test t bilatéral au seuil α rejette l'hypothèse nulle si et seulement si μ_0 n'est pas dans l'intervalle de confiance bilatéral pour μ avec niveau de confiance $100 \times (1-\alpha)\%$.

Test d'hypothèse sur la variance

- Soit un échantillon aléatoire

$$X_1, \ldots, X_n$$

de la distribution normale avec moyenne μ et variance σ^2.

- Comment peut-on tester des hypothèses sur σ^2 ?

Test d'hypothèse sur la variance

- Soit un échantillon aléatoire

$$X_1, \ldots, X_n$$

de la distribution normale avec moyenne μ et variance σ^2.

- Comment peut-on tester des hypothèses sur σ^2 ?
- On rappelle la statistique khi-carrée

$$\frac{(n-1)}{\sigma_0^2} S_n^2,$$

qui suit une distribution $\chi^2_{(n-1)}$ quand $\sigma^2 = \sigma_0^2$.

Test du khi-carré

Les problèmes sur la variance d'une population normale

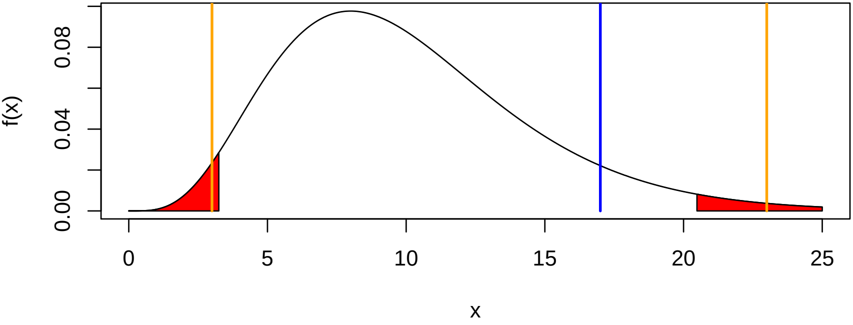
$$\begin{aligned} \text{problème I.} \quad & \mathcal{H}_0 \ : \ \sigma^2 \leq \sigma_0^2, \quad \mathcal{H}_1 \ : \ \sigma^2 > \sigma_0^2, \\ \text{problème II.} \quad & \mathcal{H}_0 \ : \ \sigma^2 \geq \sigma_0^2, \quad \mathcal{H}_1 \ : \ \sigma^2 < \sigma_0^2, \\ \text{problème III.} \quad & \mathcal{H}_0 \ : \ \sigma^2 = \sigma_0^2, \quad \mathcal{H}_1 \ : \ \sigma^2 \neq \sigma_0^2, \end{aligned}$$

peuvent être testés avec le test du khi-carré. Le test au seuil α rejette $\mathcal{H}_0$ si

$$\begin{aligned} \text{problème I.} \quad & \frac{(n-1)}{\sigma_0^2} S_n^2 \geq \chi^2_{n-1,\alpha}, \\ \text{problème II.} \quad & \frac{(n-1)}{\sigma_0^2} S_n^2 \leq \chi^2_{n-1,1-\alpha}, \\ \text{problème III.} \quad & \frac{(n-1)}{\sigma_0^2} S_n^2 \leq \chi^2_{n-1,1-\alpha/2} \quad \text{ou} \quad \frac{(n-1)}{\sigma_0} S_n^2 \geq \chi^2_{n-1,\alpha/2} \end{aligned}$$

Illustration de la région critique

Région critique (en rouge) du test khi-carré bilatéral quand $n = 11$ et $\alpha = 0.05$. La ligne bleue est une valeur observée de la statistique du test pour laquelle on ne rejette pas $\mathcal{H}_0$, les lignes oranges sont des valeurs observées de la statistique du test pour lesquelles on rejette $\mathcal{H}_0$.



Exemple

Une compagnie produit des pièces de moteur. La variance de leur diamètre doit être inférieure ou égale à 0.0002.

Un échantillon aléatoire de 10 pièces mène à une variance échantillonnale de 0.0003. Pour tester

$$\mathcal{H}_0 \; : \; \sigma^2 \leq 0.0002 \quad \text{vs.} \quad \mathcal{H}_1 \; : \; \sigma^2 > 0.0002$$

au seuil de signification $\alpha = 0.05$, la valeur critique est

$$\chi^2_{9,0.05} = 16.919.$$

Sachant que la valeur observée de la statistique du test est

$$\frac{10-1}{0.0002} \times 0.0003 = 13.5 < 16.919,$$

il n'y a pas suffisamment de preuve pour rejeter $\mathcal{H}_0$.

Exemple (suite)

Le seuil observé du test est

$$\Pr(W_{n-1} \geq 13.5),$$

où W_{n-1} est une variable aléatoire $\chi^2_{(n-1=9)}$.

```
stat <- (9/0.0002)*0.0003
1-pchisq(stat,df=9)

## [1] 0.1412558
```

Exemple

Une scientifique était convaincue que la variabilité de son équipement de mesure menait à un écart-type de 2 ; $n = 16$ mesures ont donné $s_n^2 = 6.1$. Est-ce que les données sont en désaccord avec son affirmation ?

Exemple

Une scientifique était convaincue que la variabilité de son équipement de mesure menait à un écart-type de 2 ; $n = 16$ mesures ont donné $s_n^2 = 6.1$. Est-ce que les données sont en désaccord avec son affirmation ?

Ici, le problème est

$$\mathcal{H}_0 \ : \ \sigma^2 = 4 \quad \text{vs.} \quad \mathcal{H}_1 \ : \ \sigma^2 \neq 4.$$

La valeur observée de la statistique du test est

$$\frac{16-1}{4} \times 6.1 = 22.875.$$

Les valeurs critiques du test khi-carré bilatéral au seuil $\alpha = 0.05$ sont

$$\chi^2_{15,0.025} = 6.26214, \quad \chi^2_{15,0.975} = 27.48839.$$

Comme la statistique est entre ces deux points, $\mathcal{H}_0$ ne peut pas être rejetée.

Exemple (suite)

Pour calculer le seuil observé, on doit tenir compte des deux queues de la distribution khi-carrée.

Comme la valeur observée de la statistique du test est dans la queue de droite de la distribution khi-carrée, le seuil observé est

$$2 \times \Pr(W_{n-1} \geq 22.875).$$

```
stat <- (15/4)*6.1
2*(1-pchisq(stat,df=15))

## [1] 0.1736603
```

Cela confirme que $\mathcal{H}_0$ n'est pas rejetée, même au seuil $\alpha = 0.1$.

Exemple (fin !)

Comme pour le test t, il y a une correspondance entre le test khi-carré et l'intervalle de confiance pour la variance pour des échantillons normaux.

L'intervalle de confiance bilatéral pour σ^2 avec niveau de confiance 95% est

$$\left[\frac{15}{\chi^2_{15,0.975}} \times 6.1, \frac{15}{\chi^2_{15,0.975}} \times 6.1\right] = [3.329, 14.612]$$

Comme $\sigma^2 = 4$ est dans l'intervalle, il n'y a pas assez d'information pour conclure que la variance n'est pas égale à 4, au niveau de confiance 95%.

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).
- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).
- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).
- On a développé des tests d'hypothèses sur la moyenne et la variance pour des échantillons normaux.

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).
- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).
- On a développé des tests d'hypothèses sur la moyenne et la variance pour des échantillons normaux.
- Comment peut-on tester des hypothèses complexes lorsque les données ne sont pas normales ?

Problèmes génŕaux

On veut tester l'un des trois problèmes sur un paramètre θ:

$$\begin{aligned} \text{problème I.} \quad & \mathcal{H}_0 \ : \ \theta \leq \theta_0, \quad \mathcal{H}_1 \ : \ \theta > \theta_0, \\ \text{problème II.} \quad & \mathcal{H}_0 \ : \ \theta \geq \theta_0, \quad \mathcal{H}_1 \ : \ \theta < \theta_0, \\ \text{problème III.} \quad & \mathcal{H}_0 \ : \ \theta = \theta_0, \quad \mathcal{H}_1 \ : \ \theta \neq \theta_0. \end{aligned}$$

Problèmes génŕaux

On veut tester l'un des trois problèmes sur un paramètre θ:

$$\begin{aligned} \text{problème I.} \quad & \mathcal{H}_0 : \theta \leq \theta_0, \quad \mathcal{H}_1 : \theta > \theta_0, \\ \text{problème II.} \quad & \mathcal{H}_0 : \theta \geq \theta_0, \quad \mathcal{H}_1 : \theta < \theta_0, \\ \text{problème III.} \quad & \mathcal{H}_0 : \theta = \theta_0, \quad \mathcal{H}_1 : \theta \neq \theta_0. \end{aligned}$$

Le paramètre θ est habituellement :

- un paramètre général, comme la moyenne ou la variance;
- un paramètre spécifique d'un modèle paramétrique.

Manière générale de construire un test

Pour construire un test approprié, on trouve une statistique de test T telle que :

- T est une statistique : une fonction des données et de constantes connues, souvent θ_0.

Manière générale de construire un test

Pour construire un test approprié, on trouve une statistique de test T telle que :

- T est une statistique : une fonction des données et de constantes connues, souvent θ_0.
- On s'attend à ce que T prennent des valeurs spéciales sous $\mathcal{H}_1$ ou $\mathcal{H}_0$ (négatives, positives, près de zéro, etc.).

Manière générale de construire un test

Pour construire un test approprié, on trouve une statistique de test T telle que :

- T est une statistique : une fonction des données et de constantes connues, souvent θ_0.
- On s'attend à ce que T prennent des valeurs spéciales sous $\mathcal{H}_1$ ou $\mathcal{H}_0$ (négatives, positives, près de zéro, etc.).
- On peut identifier la distribution de T sous $\mathcal{H}_0$, ou du moins lorsque $\theta = \theta_0$, au moins lorsque n est grand.

Déjà vu

On rappelle que pour plusieurs estimateurs $\hat{\theta}_n$ de θ, quand $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}} \rightsquigarrow \mathcal{N}(0, 1)$$

où $\hat{\sigma}_{\hat{\theta}_n}$ est un estimateur convergent de l'écart-type (asymptotique) de $\hat{\theta}_n$.

Déjà vu

On rappelle que pour plusieurs estimateurs $\hat{\theta}_n$ de θ, quand $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}} \rightsquigarrow \mathcal{N}(0, 1)$$

où $\hat{\sigma}_{\hat{\theta}_n}$ est un estimateur convergent de l'écart-type (asymptotique) de $\hat{\theta}_n$.

On avait utilisé ce résultat pour construire des intervalles de confiance pour θ basés sur de grands échantillons.

Exemples

Cette propriété s'applique à :

- ✓ la moyenne échantillonnale et la différence de moyennes échantillonnales d'échantillons indépendants ;

Exemples

Cette propriété s'applique à :

- ✓ la moyenne échantillonnale et la différence de moyennes échantillonnales d'échantillons indépendants ;
- ✓ la proportion échantillonnale et la différence de proportions échantillonnales dans des échantillons indépendants;

Exemples

Cette propriété s'applique à :

- ✓ la moyenne échantillonnale et la différence de moyennes échantillonnales d'échantillons indépendants ;
- ✓ la proportion échantillonnale et la différence de proportions échantillonnales dans des échantillons indépendants;
- ✓ les estimateurs du maximum de vraisemblance (sous certaines conditions);

Exemples

Cette propriété s'applique à :

- ✓ la moyenne échantillonnale et la différence de moyennes échantillonnales d'échantillons indépendants ;
- ✓ la proportion échantillonnale et la différence de proportions échantillonnales dans des échantillons indépendants;
- ✓ les estimateurs du maximum de vraisemblance (sous certaines conditions);
- ✓ plusieurs estimateurs des moments.

Statistique du test

Sous l'hypothèse que $\theta = \theta_0$, la statistique

$$Z_n = \frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_{\hat{\theta}_n}}$$

est approximativement normale centrée réduite.

Statistique du test

Sous l'hypothèse que $\theta = \theta_0$, la statistique

$$Z_n = \frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_{\hat{\theta}_n}}$$

est approximativement normale centrée réduite.

On s'attend à ce que les valeurs de Z_n soient négatives, positives et proche de zéro sous $\mathcal{H}_0$ dans les problèmes I, II et III, respectivement.

Tests z pour grands échantillons

Soient les problèmes

$$\begin{array}{rl} \text{problème I.} & \mathcal{H}_0 : \theta \leq \theta_0, \quad \mathcal{H}_1 : \theta > \theta_0, \\ \text{problème II.} & \mathcal{H}_0 : \theta \geq \theta_0, \quad \mathcal{H}_1 : \theta < \theta_0, \\ \text{problème III.} & \mathcal{H}_0 : \theta = \theta_0, \quad \mathcal{H}_1 : \theta \neq \theta_0. \end{array}$$

Test z pour grand échantillon

Soit $Z_n = (\hat{\theta}_n - \theta_0)/\hat{\sigma}_{\hat{\theta}_n}$. Le test z pour grand échantillon au seuil α rejette $\mathcal{H}_0$ si

$$\begin{array}{rl} \text{problème I.} & Z_n \geq z_\alpha, \\ \text{problème II.} & Z_n \leq -z_\alpha, \\ \text{problème III.} & |Z_n| \geq z_{\alpha/2}, \end{array}$$

où z_α est le quantile d'ordre $1-\alpha$ de la $\mathcal{N}(0,1)$.

Relations avec intervalles de confiance

Pour construire des intervalles de confiance pour θ basés sur de grands échantillons, nous avons utilisé

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}}$$

comme pivot approximatif.

Relations avec intervalles de confiance

Pour construire des intervalles de confiance pour θ basés sur de grands échantillons, nous avons utilisé

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}}$$

comme pivot approximatif.

- Le test z pour grand échantillon au seuil α de l'hypothèse $\mathcal{H}_0 : \theta \leq (\geq)\, \theta_0$ rejette $\mathcal{H}_0$ si et seulement si θ_0 n'est pas inclus dans l'intervalle de confiance unilatéral à gauche (à droite) pour θ au niveau de confiance $100 \times (1-\alpha)\%$.

Relations avec intervalles de confiance

Pour construire des intervalles de confiance pour θ basés sur de grands échantillons, nous avons utilisé

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}}$$

comme pivot approximatif.

- Le test z pour grand échantillon au seuil α de l'hypothèse $\mathcal{H}_0 : \theta \leq (\geq)\, \theta_0$ rejette $\mathcal{H}_0$ si et seulement si θ_0 n'est pas inclus dans l'intervalle de confiance unilatéral à gauche (à droite) pour θ au niveau de confiance $100 \times (1-\alpha)\%$.

- Le test z pour grand échantillon au seuil α de l'hypothèse $\mathcal{H}_0 : \theta = \theta_0$ rejette $\mathcal{H}_0$ si et seulement si θ_0 n'est pas inclus dans l'intervalle de confiance bilatéral pour θ au niveau de confiance $100 \times (1-\alpha)\%$.

Exemple : moyenne

Une étude concernant l'allergie à la poudre sur les gants de latex a été publiée en 2004 dans *Current Allergy & Clinical Immunology*. On y étudie l'exposition à la poudre de $n = 46$ employés de l'hôpital allergiques au latex. Le nombre de gants de latex utilisés par semaine par ces personnes est

$$\bar{x}_n = 19.3, \quad s_n = 11.9.$$

Exemple : moyenne

Une étude concernant l'allergie à la poudre sur les gants de latex a été publiée en 2004 dans *Current Allergy & Clinical Immunology*. On y étudie l'exposition à la poudre de $n = 46$ employés de l'hôpital allergiques au latex. Le nombre de gants de latex utilisés par semaine par ces personnes est

$$\bar{x}_n = 19.3, \quad s_n = 11.9.$$

Peut-on conclure que le nombre moyen de gants de latex utilisés par semaine par des employés de l'hôpital avec une allergie au latex est supérieur à 15 ?

Exemple : moyenne

Une étude concernant l'allergie à la poudre sur les gants de latex a été publiée en 2004 dans *Current Allergy & Clinical Immunology*. On y étudie l'exposition à la poudre de $n = 46$ employés de l'hôpital allergiques au latex. Le nombre de gants de latex utilisés par semaine par ces personnes est

$$\bar{x}_n = 19.3, \quad s_n = 11.9.$$

Peut-on conclure que le nombre moyen de gants de latex utilisés par semaine par des employés de l'hôpital avec une allergie au latex est supérieur à 15 ?

Ici, le paramètre d'intérêt est la moyenne et

$$\mathcal{H}_0 \ : \ \mu \leq 15, \quad \mathcal{H}_1 \ : \ \mu > 15.$$

Exemple (suite)

Par le Théorème central limite et le lemme de Slutsky,

$$\sqrt{n} \times \frac{\bar{X}_n - \mu}{S_n}$$

est approximativement normale centrée réduite lorsque n est grand.

Exemple (suite)

Par le Théorème central limite et le lemme de Slutsky,

$$\sqrt{n} \times \frac{\bar{X}_n - \mu}{S_n}$$

est approximativement normale centrée réduite lorsque n est grand. La statistique du test est donc

$$Z_n = \sqrt{n} \times \frac{\bar{X}_n - 15}{S_n}$$

et sa valeur observée est

$$z_n = \sqrt{46} \times \frac{19.3 - 15}{11.9} = 2.45076.$$

Exemple (suite)

Par le Théorème central limite et le lemme de Slutsky,

$$\sqrt{n} \times \frac{\bar{X}_n - \mu}{S_n}$$

est approximativement normale centrée réduite lorsque n est grand. La statistique du test est donc

$$Z_n = \sqrt{n} \times \frac{\bar{X}_n - 15}{S_n}$$

et sa valeur observée est

$$z_n = \sqrt{46} \times \frac{19.3 - 15}{11.9} = 2.45076.$$

Le test z rejette $\mathcal{H}_0$ au seuil $\alpha = 0.01$:

$$z_n = 2.45076 > 2.326 = z_{0.01}.$$

Seuil observé pour tests z

Supposons que

$$z_n = \frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_{\hat{\theta}_n}}$$

est la valeur observée de la statistique du test Z_n.

Seuil observé pour tests z

Supposons que

$$z_n = \frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_{\hat{\theta}_n}}$$

est la valeur observée de la statistique du test Z_n.

Seuil observé ou valeur p

Le seuil observé ou valeur p du test z est :

- $1 - \Phi(z_n)$ dans le problème I $\mathcal{H}_0 : \theta \leq \theta_0$,
- $\Phi(z_n)$ dans le problème II $\mathcal{H}_0 : \theta \geq \theta_0$,
- $2\{1 - \Phi(|z_n|)\}$ dans le problème III $\mathcal{H}_0 : \theta = \theta_0$,

où Φ est la fonction de répartition de la loi normale centrée réduite.

Exemple (suite)

Dans l'exemple sur l'exposition à la poudre sur les gants de latex,

$$\mathcal{H}_0 \ : \ \mu \leq 15, \quad \mathcal{H}_1 \ : \ \mu > 15$$

et la valeur observée de la statistique du test était

$$z_n = 2.45076.$$

Exemple (suite)

Dans l'exemple sur l'exposition à la poudre sur les gants de latex,

$$\mathcal{H}_0 \;:\; \mu \leq 15, \quad \mathcal{H}_1 \;:\; \mu > 15$$

et la valeur observée de la statistique du test était

$$z_n = 2.45076.$$

Le seuil observé est donc

$$1 - \Phi(2.4507579) = 0.00713,$$

ce qui confirme que $\mathcal{H}_0$ est rejetée au seuil $\alpha = 0.01$, en particulier.

Exemple: différence de moyennes

Une étude psychologique a été menée pour comparer le temps de réaction des hommes et des femmes à un stimulus. Des échantillons aléatoires indépendants de 50 hommes et 50 femmes sont utilisés dans l'expérience. Les résultats (en secondes) sont résumés plus bas :

Hommes	Femmes
$n = 50$	$m = 50$
$\bar{x}_n = 3.6$	$\bar{y}_m = 3.8$
$s_n^2 = 0.18$	$s_m^2 = 0.14$

Est-ce que les données suggèrent que le temps de réaction est différent pour les hommes et les femmes ?

Exemple: différence de moyennes

Une étude psychologique a été menée pour comparer le temps de réaction des hommes et des femmes à un stimulus. Des échantillons aléatoires indépendants de 50 hommes et 50 femmes sont utilisés dans l'expérience. Les résultats (en secondes) sont résumés plus bas :

Hommes	Femmes
$n = 50$	$m = 50$
$\bar{x}_n = 3.6$	$\bar{y}_m = 3.8$
$s_n^2 = 0.18$	$s_m^2 = 0.14$

Est-ce que les données suggèrent que le temps de réaction est différent pour les hommes et les femmes ?

Ici, le paramètre d'intérêt est $\theta = \mu_1 - \mu_2$ et le problème est

$$\mathcal{H}_0 \ : \ \mu_1 - \mu_2 = 0, \quad \mathcal{H}_1 \ : \ \mu_1 - \mu_2 \neq 0.$$

Exemple (suite)

Pour tester des hypothèses sur la différence de moyennes, une statistique de test appropriée est

$$Z_n = \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_{1,0} - \mu_{2,0})}{\sqrt{\frac{S_n^2}{n} + \frac{S_m^2}{m}}},$$

Exemple (suite)

Pour tester des hypothèses sur la différence de moyennes, une statistique de test appropriée est

$$Z_n = \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_{1,0} - \mu_{2,0})}{\sqrt{\frac{S_n^2}{n} + \frac{S_m^2}{m}}},$$

où $\mu_{1,0} - \mu_{2,0} = \theta_0$ est la différence entre les moyennes sous l'hypothèse nulle; S_n^2 et S_m^2 sont les variances échantillonnales des premier et deuxième échantillons, respectivement, et $\bar{X}_n$ et $\bar{Y}_m$ sont les moyennes échantillonnales des premier et deuxième échantillons, respectivement.

Exemple (suite)

Pour tester des hypothèses sur la différence de moyennes, une statistique de test appropriée est

$$Z_n = \frac{(\bar{X}_n - \bar{Y}_m) - (\mu_{1,0} - \mu_{2,0})}{\sqrt{\frac{S_n^2}{n} + \frac{S_m^2}{m}}},$$

où $\mu_{1,0} - \mu_{2,0} = \theta_0$ est la différence entre les moyennes sous l'hypothèse nulle; S_n^2 et S_m^2 sont les variances échantillonnales des premier et deuxième échantillons, respectivement, et $\bar{X}_n$ et $\bar{Y}_m$ sont les moyennes échantillonnales des premier et deuxième échantillons, respectivement.

Dans cet exemple, la valeur observée de la statistique du test est

$$z_n = \frac{3.6 - 3.8 - 0}{\sqrt{\frac{0.18}{50} + \frac{0.14}{50}}} = -2.5.$$

Exemple (suite)

Supposons que l'on veuille tester au seuil de signification $\alpha = 0.05$.
Puisque la contre-hypothèse est bilatérale, le quantile normal à utiliser est

$$z_{0.025} = 1.96.$$

Exemple (suite)

Supposons que l'on veuille tester au seuil de signification $\alpha = 0.05$.
Puisque la contre-hypothèse est bilatérale, le quantile normal à utiliser est

$$z_{0.025} = 1.96.$$

Comme

$$|z_n| = 2.5 > z_{0.025},$$

$\mathcal{H}_0$ est rejetée au seuil 0.05, i.e., on peut conclure que le temps moyen de réaction est différent pour les hommes et les femmes, au seuil de 5 %.

Exemple (suite)

Supposons que l'on veuille tester au seuil de signification $\alpha = 0.05$.
Puisque la contre-hypothèse est bilatérale, le quantile normal à utiliser est

$$z_{0.025} = 1.96.$$

Comme

$$|z_n| = 2.5 > z_{0.025},$$

$\mathcal{H}_0$ est rejetée au seuil 0.05, i.e., on peut conclure que le temps moyen de réaction est différent pour les hommes et les femmes, au seuil de 5 %.

Le seuil observé est

$$2\{1 - \Phi(|z_n|)\} = 2\{1 - \Phi(2.5)\} = 0.012.$$

Exemple: Test de Wald

On a ajusté un modèle paramétrique avec FMP ou densité $f(\cdot;\theta)$ à un échantillon aléatoire par la méthode du maximum de vraisemblance.

Exemple: Test de Wald

On a ajusté un modèle paramétrique avec FMP ou densité $f(\cdot;\theta)$ à un échantillon aléatoire par la méthode du maximum de vraisemblance.

Pour tester des hypothèses sur $\theta \in \mathbb{R}$, on peut utiliser le fait que, sous certaines conditions sur $f(\cdot;\theta)$, l'EMV $\hat{\theta}_n$ de θ est tel que

$$\frac{\hat{\theta}_n - \theta}{\sqrt{1/\{\widehat{nI(\theta)}\}}} \rightsquigarrow \mathcal{N}(0,1)$$

quand $n \to \infty$, où $\widehat{nI(\theta)}$ est un estimateur convergent de $nI(\theta)$, basé soit sur l'information espérée ou observée.

Exemple: Test de Wald

On a ajusté un modèle paramétrique avec FMP ou densité $f(\cdot;\theta)$ à un échantillon aléatoire par la méthode du maximum de vraisemblance.

Pour tester des hypothèses sur $\theta \in \mathbb{R}$, on peut utiliser le fait que, sous certaines conditions sur $f(\cdot;\theta)$, l'EMV $\hat{\theta}_n$ de θ est tel que

$$\frac{\hat{\theta}_n - \theta}{\sqrt{1/\{\widehat{nI(\theta)}\}}} \rightsquigarrow \mathcal{N}(0,1)$$

quand $n \to \infty$, où $\widehat{nI(\theta)}$ est un estimateur convergent de $nI(\theta)$, basé soit sur l'information espérée ou observée.

Le test z pour les hypothèses sur θ est basé sur la statistique de Wald

$$Z_n = \frac{\hat{\theta}_n - \theta_0}{\sqrt{1/\{\widehat{nI(\theta)}\}}}.$$

Exemple (suite)

On vérifie si des ressorts sont fiables en les chargeant de façon répétée, jusqu'à ce qu'ils brisent. Les “temps” de défaillance à la tension 950 N/mm^2 sont

```
springs <- c(225,171,198,189,189,135,162,135,117,162)
```

Exemple (suite)

On vérifie si des ressorts sont fiables en les chargeant de façon répétée, jusqu'à ce qu'ils brisent. Les "temps" de défaillance à la tension 950 N/mm^2 sont

```r
springs <- c(225,171,198,189,189,135,162,135,117,162)
```

On ajuste une distribution exponentielle à ces données et on veut tester

$$\mathcal{H}_0 : \theta \leq 150, \quad \mathcal{H}_1 : \theta > 150.$$

Exemple (suite)

On vérifie si des ressorts sont fiables en les chargeant de façon répétée, jusqu'à ce qu'ils brisent. Les "temps" de défaillance à la tension 950 N/mm^2 sont

```
springs <- c(225,171,198,189,189,135,162,135,117,162)
```

On ajuste une distribution exponentielle à ces données et on veut tester

$$\mathcal{H}_0 \,:\, \theta \leq 150, \quad \mathcal{H}_1 \,:\, \theta > 150.$$

L'EMV de θ et son écart-type estimé sont

$$\hat{\theta}_n = \bar{x}_n = 168.3, \quad \sqrt{1/\widehat{nI(\theta)}} = \frac{\bar{x}_n}{\sqrt{n}} = 53.22.$$

Exemple (suite)

On vérifie si des ressorts sont fiables en les chargeant de façon répétée, jusqu'à ce qu'ils brisent. Les "temps" de défaillance à la tension 950 N/mm^2 sont

```
springs <- c(225,171,198,189,189,135,162,135,117,162)
```

On ajuste une distribution exponentielle à ces données et on veut tester

$$\mathcal{H}_0 \,:\, \theta \leq 150, \quad \mathcal{H}_1 \,:\, \theta > 150.$$

L'EMV de θ et son écart-type estimé sont

$$\hat{\theta}_n = \bar{x}_n = 168.3, \quad \sqrt{1/\widehat{nI(\theta)}} = \frac{\bar{x}_n}{\sqrt{n}} = 53.22.$$

La valeur observée de la statistique de Wald est donc

$$z_n = \frac{168.3 - 150}{53.22} = 0.34385.$$

Exemple (suite)

On vérifie si des ressorts sont fiables en les chargeant de façon répétée, jusqu'à ce qu'ils brisent. Les "temps" de défaillance à la tension 950 N/mm^2 sont

```
springs <- c(225,171,198,189,189,135,162,135,117,162)
```

On ajuste une distribution exponentielle à ces données et on veut tester

$$\mathcal{H}_0 \,:\, \theta \leq 150, \quad \mathcal{H}_1 \,:\, \theta > 150.$$

L'EMV de θ et son écart-type estimé sont

$$\hat{\theta}_n = \bar{x}_n = 168.3, \quad \sqrt{1/\widehat{nI(\theta)}} = \frac{\bar{x}_n}{\sqrt{n}} = 53.22.$$

La valeur observée de la statistique de Wald est donc

$$z_n = \frac{168.3 - 150}{53.22} = 0.34385.$$

Le seuil observé est égal à $1 - \Phi(z_n) = 0.37$ et $\mathcal{H}_0$ n'est pas rejetée.

Taille du test z

Dans le problème I,

$$\mathcal{H}_0 \ : \ \theta \leq \theta_0, \quad \mathcal{H}_1 \ : \ \theta > \theta_0,$$

Taille du test z

Dans le problème I,

$$\mathcal{H}_0 \;:\; \theta \leq \theta_0, \quad \mathcal{H}_1 \;:\; \theta > \theta_0,$$

la taille du test est approximativement α, parce que pour tout $\theta_1 \leq \theta_0$,

$$\begin{aligned}
\Pr\left(Z_n \geq z_\alpha | \theta = \theta_1\right) &= \Pr\left(\frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_{\hat{\theta}_n}} \geq z_\alpha \middle| \theta = \theta_1\right) \\
&= \Pr\left(\underbrace{\frac{\hat{\theta}_n - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}}}_{\approx \mathcal{N}(0,1)} \geq z_\alpha + \frac{\theta_0 - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}} \middle| \theta = \theta_1\right) \\
&\approx 1 - \Phi\left(z_\alpha + \underbrace{\frac{\theta_0 - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}}}_{\geq 0}\right) \leq \alpha.
\end{aligned}$$

Illustration I

Pourcentage de rejet de $\mathcal{H}_0 : \mu \leq 1$ avec un test z au seuil $\alpha = 0.05$ dans 1000 échantillons de différentes tailles tirés de la loi Poisson avec paramètre $\lambda = 1$. La ligne rouge est le seuil de signification désiré.

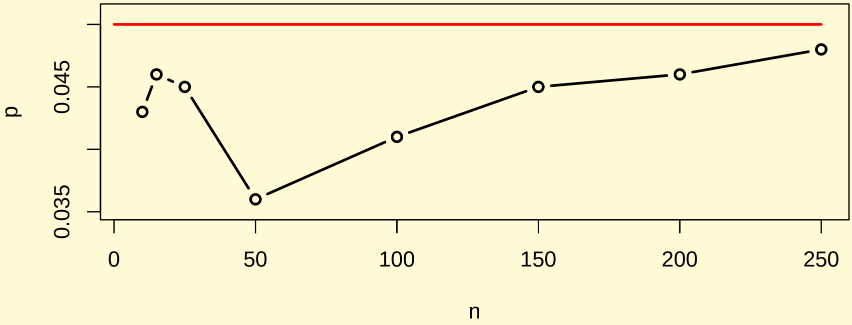
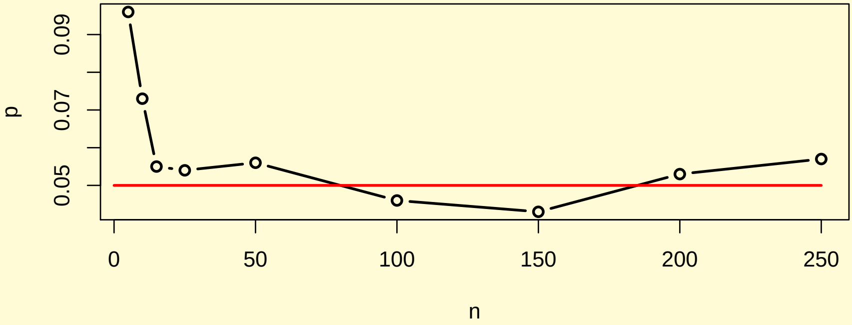


Illustration II

Pourcentage de rejet de $\mathcal{H}_0 : \mu \leq 1$ avec un test z au seuil $\alpha = 0.05$ dans 1000 échantillons de différentes tailles tirés de la loi normale avec moyenne $\mu = 1$ et variance 1. La ligne rouge est le seuil de signification désiré.



Exemple : proportion

Une machine d'usine produit 10 % d'items défectueux par jour. Un échantillon aléatoire de 100 items pour une journée contient 15 défectueux, et le superviseur affirme que la machine doit être réparée. Est-ce vrai que la machine produit plus de 10 % d'items défectueux, en moyenne ?

Exemple : proportion

Une machine d'usine produit 10 % d'items défectueux par jour. Un échantillon aléatoire de 100 items pour une journée contient 15 défectueux, et le superviseur affirme que la machine doit être réparée. Est-ce vrai que la machine produit plus de 10 % d'items défectueux, en moyenne ?

Ici, le paramètre d'intérêt est la proportion, $\theta = p$, et le problème est

$$\mathcal{H}_0 \, : \, p \leq 0.1, \quad \mathcal{H}_1 \, : \, p > 0.1.$$

On veut tester au seuil de signification $\alpha = 0.05$.

Exemple (suite)

On a vu que, quand $n \to \infty$, la proportion échantillonnale est telle que

$$\frac{p_n - p}{\sqrt{p(1-p)/n}} \rightsquigarrow \mathcal{N}(0, 1).$$

Exemple (suite)

On a vu que, quand $n \to \infty$, la proportion échantillonnale est telle que

$$\frac{p_n - p}{\sqrt{p(1-p)/n}} \rightsquigarrow \mathcal{N}(0,1).$$

Pour obtenir une statistique de test, il faut estimer la variance de p_n. On a besoin d'une statistique dont la distribution est approximativement normale centrée réduite quand $p = p_0 = 0.1$, donc une statistique de test appropriée est

$$Z_n = \frac{p_n - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{p_n - 0.1}{\sqrt{(0.1 \times 0.9)/n}}.$$

Exemple (suite)

On a vu que, quand $n \to \infty$, la proportion échantillonnale est telle que

$$\frac{p_n - p}{\sqrt{p(1-p)/n}} \rightsquigarrow \mathcal{N}(0,1).$$

Pour obtenir une statistique de test, il faut estimer la variance de p_n. On a besoin d'une statistique dont la distribution est approximativement normale centrée réduite quand $p = p_0 = 0.1$, donc une statistique de test appropriée est

$$Z_n = \frac{p_n - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{p_n - 0.1}{\sqrt{(0.1 \times 0.9)/n}}.$$

Sa valeur observée z_n est

$$z_n = \frac{0.15 - 0.1}{\sqrt{(0.1 \times 0.9)/100}} = 1.667.$$

Exemple (suite)

Le quantile de la distribution normale à utiliser est

$$z_{0.05} = 1.645.$$

Puisque cette valeur est plus petite que z_n, $\mathcal{H}_0$ est rejetée au seuil $\alpha = 0.05$.

Exemple (suite)

Le quantile de la distribution normale à utiliser est

$$z_{0.05} = 1.645.$$

Puisque cette valeur est plus petite que z_n, $\mathcal{H}_0$ est rejetée au seuil $\alpha = 0.05$.

Le seuil observé est

$$1 - \Phi(z_n) = 1 - \Phi(1.667) = 0.048.$$

Donc on ne pourrait pas rejeter $\mathcal{H}_0$ su seuil de signification 0.01, par exemple.

Puissance du test z

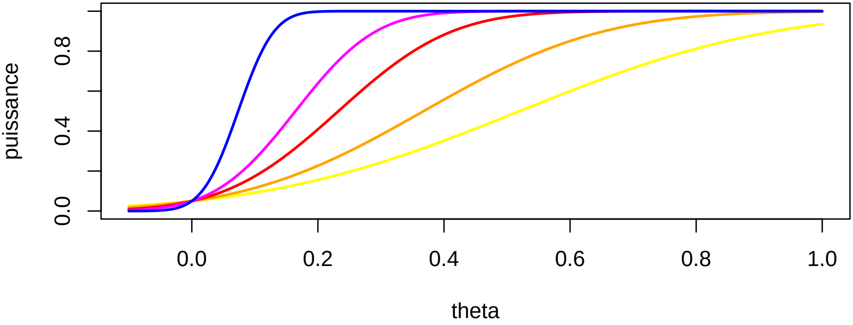
Pour le problème I et pour tout $\theta_1 > \theta_0$,

$$\begin{aligned}
\pi(\theta_1) &= \Pr\left(\frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_{\hat{\theta}_n}} \geq z_\alpha \middle| \theta = \theta_1\right) \\
&= \Pr\left(\underbrace{\frac{\hat{\theta}_n - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}}}_{\approx \mathcal{N}(0,1)} \geq z_\alpha + \frac{\theta_0 - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}} \middle| \theta = \theta_1\right) \\
&\approx 1 - \Phi\left(z_\alpha + \underbrace{\frac{\theta_0 - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}}}_{<0}\right) > \alpha.
\end{aligned}$$

La quantité orange est négative et décroissante en θ_1, et donc la puissance du test est croissante en θ_1.

Illustration

Puissance du test z unilatéral avec $\mathcal{H}_0 : \mu \leq \theta_0$, $\theta_0 = 0$, $\alpha = 0.05$ et $n = 10$, $n = 20$, $n = 50$, $n = 100$, et $n = 500$. Pour fins d'illustration, $\hat{\sigma}_{\hat{\theta}_n} = 1/\sqrt{n}$.



Erreur de type II

Lorsque $\theta_1 > \theta_0$, c'est-à-dire quand la contre-hypothèse est vraie,

$$\Pr(\text{Accepter } \mathcal{H}_0 | \theta = \theta_1) = 1 - \pi(\theta_1).$$

Cela veut dire que l'erreur de type II quand $\theta = \theta_1$ est donnée par

$$1 - \pi(\theta_1) \approx \Phi\left(z_\alpha + \frac{\theta_0 - \theta_1}{\hat{\sigma}_{\hat{\theta}_n}} \right).$$

- L'erreur de type II diminue lorsque la distance entre θ_1 et θ_0 augmente.
- Puisque $\sigma_{\hat{\theta}_n}$ est souvent de la forme $\sigma/\sqrt{n}$, l'erreur de type II diminue quand la taille d'échantillon augmente.

Exemple

Le vice-président d'une compagnie veut être capable de détecter une différence d'un appel dans le nombre moyen d'appels de client par semaine. Pour faire cela, il choisit aléatoirement $n = 36$ vendeurs, et note le nombre d'appels faits pendant une semaine sélectionnée aléatoirement. La moyenne échantillonnale et la variance échantillonnale sont $\bar{x}_n = 15.5$ et $s_n^2 = 9$.
On veut tester au seuil de signification $\alpha = 0.05$

$$\mathcal{H}_0 \, : \, \mu \leq 15, \quad \mathcal{H}_1 \, : \, \mu > 15.$$

La valeur observée de la statistique du test est

$$z_n = \sqrt{n} \times \frac{\bar{x}_n - 15}{s_n} = \sqrt{36} \times \frac{15.5 - 15}{3} = 1.$$

et sachant que $z_{0.05} = 1.645$, $\mathcal{H}_0$ ne peut pas être rejetée.

Exemple (suite)

La probabilité approximative d'erreur de type II quand $\mu = 16$ est

$$\Phi\left(z_\alpha + \frac{\mu_0 - \mu_1}{s_n/\sqrt{n}}\right) = \Phi\left(z_{0.05} + \sqrt{36} \times \frac{15-16}{3}\right)$$
$$= \Phi(1.645 - 2)$$
$$= \Phi(-0.355) = 0.361.$$

Cela veut dire que dans un échantillons de taille $n = 36$, une différence de 1 unité de $\mu_0 = 15$ ne sera pas souvent détectée.

Exemple (suite)

Combien de vendeurs devrait-il suivre pour que l'erreur de type II soit β ? Il faut que

$$\Phi\left(z_\alpha + \frac{\mu_0 - \mu_1}{s_n/\sqrt{n}}\right) = \beta,$$

ce qui est équivalent à

$$z_\alpha + \sqrt{n} \times \frac{\mu_0 - \mu_1}{s_n} = -z_\beta \quad \Leftrightarrow \quad n = \frac{(z_\alpha + z_\beta)^2 s_n^2}{(\mu_1 - \mu_0)^2}.$$

Dans cet exemple, si on veut $\beta = 0.05$, on trouve

$$n = \frac{9(1.645 + 1.645)^2}{(16 - 15)^2} = 97.4169.$$

Donc, au moins $n = 98$ vendeurs devraient être sélectionnés.

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).
- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).

- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).

- On a développé des tests d'hypothèses sur la moyenne et la variance pour des échantillons normaux.

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).
- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).
- On a développé des tests d'hypothèses sur la moyenne et la variance pour des échantillons normaux.
- On a développé des tests pour des grands échantillons sur des hypothèses concernant un seul paramètre.

Où en sommes-nous ?

- Nous avons discuté des principes de test d'hypothèse (erreurs de types I et II, taille, puissance, etc.).
- On a vu la théorie concernant les tests optimaux dans des cas spéciaux (hypothèses simples, test unilatéral sur la moyenne pour un échantillon normal avec variance connue).
- On a développé des tests d'hypothèses sur la moyenne et la variance pour des échantillons normaux.
- On a développé des tests pour des grands échantillons sur des hypothèses concernant un seul paramètre.
- Comment peut-on tester des hypothèses plus complexes concernant plusieurs paramètres ?

Un second regard sur le Lemme de Neyman–Pearson

Supposons que l'échantillon aléatoire étudié

$$X_1, \ldots, X_n$$

est tiré d'une distribution paramétrique avec FMP ou densité $f(x;\theta)$, $\theta \in \{\theta_0, \theta_1\}$.

Un second regard sur le Lemme de Neyman–Pearson

Supposons que l'échantillon aléatoire étudié

$$X_1, \ldots, X_n$$

est tiré d'une distribution paramétrique avec FMP ou densité $f(x;\theta)$, $\theta \in \{\theta_0, \theta_1\}$.

Le test optimal pour le problème avec des hypothèses simples,

$$\mathcal{H}_0 \,:\, \theta = \theta_0, \quad \mathcal{H}_1 \,:\, \theta = \theta_1$$

rejette $\mathcal{H}_0$ si

$$\frac{f(\boldsymbol{x};\theta_0)}{f(\boldsymbol{x};\theta_1)} < k$$

pour une valeur critique k choisie soigneusement.

Un second regard sur le Lemme de Neyman–Pearson

Supposons que l'échantillon aléatoire étudié

$$X_1, \ldots, X_n$$

est tiré d'une distribution paramétrique avec FMP ou densité $f(x;\theta)$, $\theta \in \{\theta_0, \theta_1\}$.

Le test optimal pour le problème avec des hypothèses simples,

$$\mathcal{H}_0 \; : \; \theta = \theta_0, \quad \mathcal{H}_1 \; : \; \theta = \theta_1$$

rejette $\mathcal{H}_0$ si

$$\frac{f(\boldsymbol{x};\theta_0)}{f(\boldsymbol{x};\theta_1)} < k$$

pour une valeur critique k choisie soigneusement. Quand $k < 1$, on peut réécrire cela comme

$$\frac{\max_{\theta=\theta_0} f(\boldsymbol{x};\theta)}{\max_{\theta\in\{\theta_0,\theta_1\}} f(\boldsymbol{x};\theta)} = \frac{\max_{\theta=\theta_0} L(\theta)}{\max_{\theta\in\{\theta_0,\theta_1\}} L(\theta)} < k.$$

Statistique du rapport de vraisemblance

On considère un échantillon aléatoire d'une distribution avec FMP ou densité $f(x;\theta)$, $\theta \in \Theta$

$$X_1, \ldots, X_n.$$

On crée une partition de Θ en Θ_0 et Θ_1 et on veut tester

$$\mathcal{H}_0 : \theta \in \Theta_0, \quad \mathcal{H}_1 : \theta \in \Theta_1.$$

Statistique du rapport de vraisemblance

On considère un échantillon aléatoire d'une distribution avec FMP ou densité $f(x;\theta)$, $\theta \in \Theta$

$$X_1, \ldots, X_n.$$

On crée une partition de Θ en Θ_0 et Θ_1 et on veut tester

$$\mathcal{H}_0 \ : \ \theta \in \Theta_0, \quad \mathcal{H}_1 \ : \ \theta \in \Theta_1.$$

Statistique du rapport de vraisemblance

La statistique du rapport de vraisemblance pour le problème ci-haut est donnée par

$$L_n = \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)},$$

où $L(\theta)$ dénote la fonction de vraisemblance pour θ.

Première observation

Pour commencer, on note que, comme

$$\Theta_0 \subset \Theta$$

il est toujours vrai que

$$0 \leq \max_{\theta \in \Theta_0} L(\theta) \leq \max_{\theta \in \Theta} L(\theta),$$

parce que le maximum sur le côté droit est pris sur un plus grand ensemble.

Première observation

Pour commencer, on note que, comme

$$\Theta_0 \subset \Theta$$

il est toujours vrai que

$$0 \leq \max_{\theta \in \Theta_0} L(\theta) \leq \max_{\theta \in \Theta} L(\theta),$$

parce que le maximum sur le côté droit est pris sur un plus grand ensemble.

Conséquemment, la statistique du rapport de vraisemblance est toujours telle que

$$0 \leq L_n \leq 1.$$

On s'attend à ce que L_n soit petit quand $\mathcal{H}_1$ est vraie et près de 1 quand $\mathcal{H}_0$ est vraie.

Exemple

Soit $X_1, \ldots, X_n$ un échantillon aléatoire de la distribution normale avec moyenne μ inconnue et variance σ^2 inconnue.

Exemple

Soit $X_1, \ldots, X_n$ un échantillon aléatoire de la distribution normale avec moyenne μ inconnue et variance σ^2 inconnue.

Pour tester

$$\mathcal{H}_0 \,:\, \mu \leq \mu_0, \quad \mathcal{H}_1 \,:\, \mu > \mu_0.$$

on peut utiliser le Test t unilatéral.

Exemple

Soit $X_1, \ldots, X_n$ un échantillon aléatoire de la distribution normale avec moyenne μ inconnue et variance σ^2 inconnue.

Pour tester

$$\mathcal{H}_0 \; : \; \mu \leq \mu_0, \quad \mathcal{H}_1 \; : \; \mu > \mu_0.$$

on peut utiliser le Test t unilatéral.

Nous verrons ici une alternative : un test basé sur la statistique du rapport de vraisemblance. Ce test rejettera $\mathcal{H}_0$ si, pour une valeur critique k choisie de façon appropriée,

$$L_n < k.$$

Exemple (suite)

On rappelle que la vraisemblance et la log-vraisemblance sont données par

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left\{-\sum_{i=1}^n \frac{(x_i - \mu)^2}{2\sigma^2}\right\},$$

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2.$$

Exemple (suite)

On rappelle que la vraisemblance et la log-vraisemblance sont données par

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left\{-\sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}\right\},$$

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

De plus,

$$\Theta = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}.$$

Exemple (suite)

On rappelle que la vraisemblance et la log-vraisemblance sont données par

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left\{-\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}\right\},$$

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2.$$

De plus,

$$\Theta = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}.$$

On a déjà montré que L (et ℓ) sont maximisées sur Θ au point

$$\hat{\mu}_n = \bar{x}_n, \quad \hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x}_n)^2.$$

Exemple (suite)

Pour calculer le maximum de ℓ sur l'espace des paramètres restreint Θ_0,

$$\Theta_0 = \{(\mu, \sigma^2) : \mu \leq \mu_0, \sigma^2 > 0\},$$

on observe que

$$\frac{\partial}{\partial \sigma^2} \ell(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2.$$

Exemple (suite)

Pour calculer le maximum de ℓ sur l'espace des paramètres restreint Θ_0,

$$\Theta_0 = \{(\mu, \sigma^2) : \mu \leq \mu_0, \sigma^2 > 0\},$$

on observe que

$$\frac{\partial}{\partial \sigma^2} \ell(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2.$$

Pour tout μ donné, $\ell(\mu, \sigma^2)$ est maximisé si

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2.$$

Exemple (suite)

$$\frac{\partial}{\partial \mu} \ell(\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = \frac{n}{\sigma^2} (\bar{x}_n - \mu).$$

Donc,

$$\frac{\partial}{\partial \mu} \ell(\mu, \sigma^2) \begin{cases} < 0 & \text{if} \quad \mu > \bar{x}_n, \\ = 0 & \text{if} \quad \mu = \bar{x}_n, \\ > 0 & \text{if} \quad \mu < \bar{x}_n \end{cases}$$

Exemple (suite)

$$\frac{\partial}{\partial \mu} \ell(\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = \frac{n}{\sigma^2} (\bar{x}_n - \mu).$$

Donc,

$$\frac{\partial}{\partial \mu} \ell(\mu, \sigma^2) \begin{cases} < 0 & \text{if} \quad \mu > \bar{x}_n, \\ = 0 & \text{if} \quad \mu = \bar{x}_n, \\ > 0 & \text{if} \quad \mu < \bar{x}_n \end{cases}$$

et $\ell(\mu, \sigma^2)$ est maximisée sur Θ_0 au point

$$\tilde{\mu}_n = \begin{cases} \mu_0 & \text{if} \quad \bar{x}_n > \mu_0, \\ \bar{x}_n & \text{if} \quad \bar{x}_n \leq \mu_0, \end{cases} \qquad \tilde{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \tilde{\mu}_n)^2.$$

Exemple (suite)

Pour calculer la statistique du rapport de vraisemblance, on note que

$$\sum_{i=1}^{n} \frac{(x_i - \hat{\mu}_n)^2}{2\hat{\sigma}_n^2} = \sum_{i=1}^{n} \frac{(x_i - \tilde{\mu}_n)^2}{2\tilde{\sigma}_n^2} = \frac{n}{2}.$$

Exemple (suite)

Pour calculer la statistique du rapport de vraisemblance, on note que

$$\sum_{i=1}^{n} \frac{(x_i - \hat{\mu}_n)^2}{2\hat{\sigma}_n^2} = \sum_{i=1}^{n} \frac{(x_i - \tilde{\mu}_n)^2}{2\tilde{\sigma}_n^2} = \frac{n}{2}.$$

Ensuite,

$$\hat{\sigma}_n^2 = \frac{n-1}{n} s_n^2.$$

Exemple (suite)

Pour calculer la statistique du rapport de vraisemblance, on note que

$$\sum_{i=1}^{n} \frac{(x_i - \hat{\mu}_n)^2}{2\hat{\sigma}_n^2} = \sum_{i=1}^{n} \frac{(x_i - \tilde{\mu}_n)^2}{2\tilde{\sigma}_n^2} = \frac{n}{2}.$$

Ensuite,

$$\hat{\sigma}_n^2 = \frac{n-1}{n} s_n^2.$$

Finalement, si $\bar{x}_n \leq \mu_0$, $\tilde{\sigma}_n^2 = \hat{\sigma}_n^2$ alors que si $\bar{x}_n > \mu_0$,

$$\tilde{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_0)^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 + (\bar{x}_n - \mu_0)^2 = \frac{n-1}{n} s_n^2 + (\bar{x}_n - \mu_0)^2.$$

Exemple (suite)

On trouve donc que la statistique du rapport de vraisemblance est égale à

$$L_n = \frac{\max_{(\mu,\sigma^2)\in\Theta_0} L(\mu,\sigma^2)}{\max_{(\mu,\sigma^2)\in\Theta} L(\mu,\sigma^2)} = \frac{L(\tilde{\mu}_n,\tilde{\sigma}_n^2)}{L(\hat{\mu}_n,\hat{\sigma}_n^2)} = \left(\frac{\hat{\sigma}_n^2}{\tilde{\sigma}_n^2}\right)^{n/2} .$$

Si $\bar{x}_n \leq \mu_0$, ça se simplifie à 1.

Exemple (suite)

On trouve donc que la statistique du rapport de vraisemblance est égale à

$$L_n = \frac{\max_{(\mu,\sigma^2)\in\Theta_0} L(\mu,\sigma^2)}{\max_{(\mu,\sigma^2)\in\Theta} L(\mu,\sigma^2)} = \frac{L(\tilde{\mu}_n, \tilde{\sigma}_n^2)}{L(\hat{\mu}_n, \hat{\sigma}_n^2)} = \left(\frac{\hat{\sigma}_n^2}{\tilde{\sigma}_n^2}\right)^{n/2} .$$

Si $\bar{x}_n \leq \mu_0$, ça se simplifie à 1. Sinon, quand $\bar{x}_n > \mu_0$,

$$\left(\frac{\hat{\sigma}_n^2}{\tilde{\sigma}_n^2}\right)^{n/2} = \left\{\frac{1}{1 + \frac{n}{n-1}\frac{(\bar{x}_n-\mu_0)^2}{s_n^2}}\right\}^{n/2} = \left(\frac{1}{1 + \frac{1}{n-1} t_n^2}\right)^{n/2},$$

où t_n est la statistique de Student t.

Exemple (suite)

On trouve donc que la statistique du rapport de vraisemblance est égale à

$$L_n = \frac{\max_{(\mu,\sigma^2)\in\Theta_0} L(\mu,\sigma^2)}{\max_{(\mu,\sigma^2)\in\Theta} L(\mu,\sigma^2)} = \frac{L(\tilde{\mu}_n, \tilde{\sigma}_n^2)}{L(\hat{\mu}_n, \hat{\sigma}_n^2)} = \left(\frac{\hat{\sigma}_n^2}{\tilde{\sigma}_n^2}\right)^{n/2}.$$

Si $\bar{x}_n \leq \mu_0$, ça se simplifie à 1. Sinon, quand $\bar{x}_n > \mu_0$,

$$\left(\frac{\hat{\sigma}_n^2}{\tilde{\sigma}_n^2}\right)^{n/2} = \left\{\frac{1}{1+\frac{n}{n-1}\frac{(\bar{x}_n-\mu_0)^2}{s_n^2}}\right\}^{n/2} = \left(\frac{1}{1+\frac{1}{n-1}t_n^2}\right)^{n/2},$$

où t_n est la statistique de Student t.

En conclusion, L_n est proche de 0 si et seulement si t_n est grande, i.e., un test basé sur L_n mène au test t.

Une autre statistique du rapport de vraisemblance

Plutôt que d'utiliser la statistique du rapport de vraisemblance L_n, il est souvent plus simple de travailler avec

$$W_n = -2\ln(L_n) = -2\left\{\max_{\theta\in\Theta_0}\ell(\theta) - \max_{\theta\in\Theta}\ell(\theta)\right\} = 2\left\{\max_{\theta\in\Theta}\ell(\theta) - \max_{\theta\in\Theta_0}\ell(\theta)\right\}.$$

Une autre statistique du rapport de vraisemblance

Plutôt que d'utiliser la statistique du rapport de vraisemblance L_n, il est souvent plus simple de travailler avec

$$W_n = -2\ln(L_n) = -2\left\{\max_{\theta\in\Theta_0} \ell(\theta) - \max_{\theta\in\Theta} \ell(\theta)\right\} = 2\left\{\max_{\theta\in\Theta} \ell(\theta) - \max_{\theta\in\Theta_0} \ell(\theta)\right\}.$$

La statistique W_n est aussi appelée la statistique du rapport de vraisemblance.

Une autre statistique du rapport de vraisemblance

Plutôt que d'utiliser la statistique du rapport de vraisemblance L_n, il est souvent plus simple de travailler avec

$$W_n = -2\ln(L_n) = -2\left\{\max_{\theta\in\Theta_0}\ell(\theta) - \max_{\theta\in\Theta}\ell(\theta)\right\} = 2\left\{\max_{\theta\in\Theta}\ell(\theta) - \max_{\theta\in\Theta_0}\ell(\theta)\right\}.$$

La statistique W_n est aussi appelée la statistique du rapport de vraisemblance.

Comme $0 \leq L_n \leq 1$, on a

$$0 \leq W_n < \infty.$$

Aussi, W_n est grande quand $\mathcal{H}_1$ est vraie et proche de 0 quand $\mathcal{H}_0$ est vraie.

Exemple (suite)

Dans l'exemple précédent,

$$W_n = n \ln\left(1 + \frac{1}{n-1}\, t_n^2\right) \approx \frac{n}{n-1}\, t_n^2$$

pour $t_n/(n-1)$ proche de 0 (en utilisant le fait que $\ln(1+x) \approx x$).

Exemple (suite)

Dans l'exemple précédent,

$$W_n = n \ln \left(1 + \frac{1}{n-1}\, t_n^2 \right) \approx \frac{n}{n-1}\, t_n^2$$

pour $t_n/(n-1)$ proche de 0 (en utilisant le fait que $\ln(1+x) \approx x$).

Intuition : Comme la loi t avec $n-1$ degrés de liberté converge vers une $\mathcal{N}(0,1)$ quand $n \to \infty$, la distribution de W_n quand $\mu = \mu_0$ sera approximativement $\chi^2_{(1)}$ quand n est grand.

Test du rapport de vraisemblance

Théorème 10.2

Soit $\theta \in \Theta$ un paramètre de dimension p, et $\theta \in \Theta_0$ est de la forme

$$\theta = (\theta_0, \theta^*),$$

où θ_0 est fixé et θ^* est un paramètre de dimension r, avec $r < p$. Alors, sous $\mathcal{H}_0$ et pour n grand,

$$W_n = 2\left\{\max_{\theta\in\Theta} \ell(\theta) - \max_{\theta\in\Theta_0} \ell(\theta)\right\}$$

suit approximativement une distribution $\chi^2_{(p-r)}$.

Test du rapport de vraisemblance

Théorème 10.2

Soit $\theta \in \Theta$ un paramètre de dimension p, et $\theta \in \Theta_0$ est de la forme

$$\theta = (\theta_0, \theta^*),$$

où θ_0 est fixé et θ^* est un paramètre de dimension r, avec $r < p$. Alors, sous $\mathcal{H}_0$ et pour n grand,

$$W_n = 2\left\{\max_{\theta \in \Theta} \ell(\theta) - \max_{\theta \in \Theta_0} \ell(\theta)\right\}$$

suit approximativement une distribution $\chi^2_{(p-r)}$. Le test du rapport de vraisemblance rejette $\mathcal{H}_0$ si

$$W_n \geq \chi^2_{p-r,\alpha},$$

où $\chi^2_{p-r,\alpha}$ est le quantile d'ordre $1-\alpha$ de la $\chi^2_{(p-r)}$.

Exemple

Rappelons les durées d'acouchement pour 95 femmes au John Radcliffe Hospital, Oxford.

```
library(SMPracticals)
data(births)
birth.time <- births$time
```

Exemple

Rappelons les durées d'acouchement pour 95 femmes au John Radcliffe Hospital, Oxford.

```
library(SMPracticals)
data(births)
birth.time <- births$time
```

On ajuste deux distributions aux données, l'exponentielle et la Gamma :

```
library(MASS)
m1 <- fitdistr(birth.time, densfun="exponential", lower=0)
m2 <- fitdistr(birth.time, densfun="gamma", lower=c(0,0))
```

Quel modèle est le plus approprié ?

Exemple (suite)

La distribution exponentielle est un cas particulier de la distribution Gamma quand $\alpha = 1$. Autrement dit, les modèles sont emboîtés (*nested*).

Exemple (suite)

La distribution exponentielle est un cas particulier de la distribution Gamma quand $\alpha = 1$. Autrement dit, les modèles sont emboîtés (*nested*).

Pour voir si le modèle exponentiel est une simplification adéquate du modèle Gamma, on peut tester

$$\mathcal{H}_0 \,:\, \alpha = 1, \quad \mathcal{H}_1 \,:\, \alpha \neq 1$$

avec le test du rapport de vraisemblance.

Exemple (suite)

La distribution exponentielle est un cas particulier de la distribution Gamma quand $\alpha = 1$. Autrement dit, les modèles sont emboîtés (*nested*).

Pour voir si le modèle exponentiel est une simplification adéquate du modèle Gamma, on peut tester

$$\mathcal{H}_0 : \alpha = 1, \quad \mathcal{H}_1 : \alpha \neq 1$$

avec le test du rapport de vraisemblance. La statistique du test est calculée comme suit :

```
2*(m2$loglik-m1$loglik)

## [1] 76.16788
```

Exemple (suite)

La distribution exponentielle est un cas particulier de la distribution Gamma quand $\alpha = 1$. Autrement dit, les modèles sont emboîtés (*nested*).

Pour voir si le modèle exponentiel est une simplification adéquate du modèle Gamma, on peut tester

$$\mathcal{H}_0 \,:\, \alpha = 1, \quad \mathcal{H}_1 \,:\, \alpha \neq 1$$

avec le test du rapport de vraisemblance. La statistique du test est calculée comme suit :

```
2*(m2$loglik-m1$loglik)

## [1] 76.16788
```

Sous $\mathcal{H}_0$, la statistique du rapport de vraisemblance est approximativement $\chi^2_{(\nu)}$, où $\nu = 2 - 1 = 1$. Le quantile 99% est 6.635, donc $\mathcal{H}_0$ est facilement rejetée au seuil 1%.

Remarque

Précédemment, on a utilisé le critère d'information d'Akaike (AIC) pour choisir le modèle avec le meilleur ajustement.

Remarque

Précédemment, on a utilisé le critère d'information d'Akaike (AIC) pour choisir le modèle avec le meilleur ajustement.

Dans l'exemple précédent,

$$\mathrm{AIC}(\mathtt{m1}) = 2\{-\ell(\hat{\beta}) + 1\}$$

et

$$\mathrm{AIC}(\mathtt{m2}) = 2\{-\ell(\hat{\alpha}, \hat{\beta}) + 2\}$$

Remarque

Précédemment, on a utilisé le critère d'information d'Akaike (AIC) pour choisir le modèle avec le meilleur ajustement.

Dans l'exemple précédent,

$$\mathrm{AIC}(\mathtt{m1}) = 2\{-\ell(\hat{\beta}) + 1\}$$

et

$$\mathrm{AIC}(\mathtt{m2}) = 2\{-\ell(\hat{\alpha}, \hat{\beta}) + 2\}$$

Donc, le modèle m2 est choisi selon l'AIC si $\mathrm{AIC}(\mathtt{m2}) < \mathrm{AIC}(\mathtt{m1})$, ce qui équivaut à

$$2\{-\ell(\hat{\alpha}, \hat{\beta}) + 2\} < 2\{-\ell(\hat{\beta}) + 1\}$$

$$\underbrace{2}_{\text{deux fois l'espérance de } \chi^2_{(1)}} < \underbrace{2\{\ell(\hat{\alpha}, \hat{\beta}) - \ell(\hat{\beta})\}}_{\text{statistique du rapport de vraisemblance}}$$

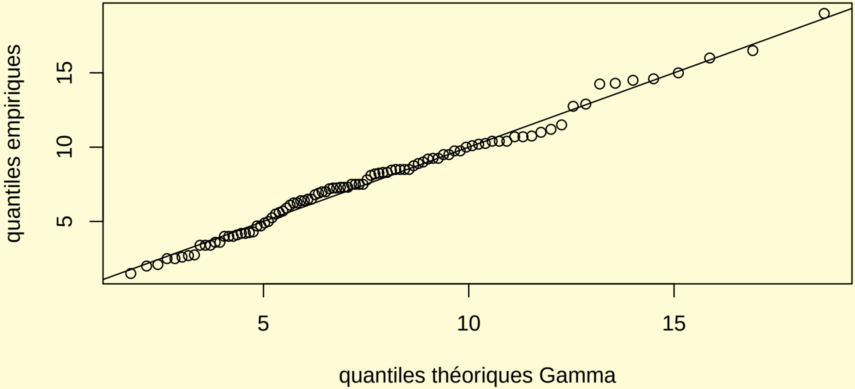
Exemple (suite)

Dans l'exemple des durées d'accouchement, on a rejeté l'hypothèse que l'exponentielle est une bonne simplification du modèle Gamma.

Mais est-ce que le modèle Gamma est adéquat pour ces données ?

Tel que vu précédemment, on peut regarder le diagramme quantile-quantile.

Exemple (suite)



Pouvons-nous faire un test formel ?

Test d'adéquation

Soit l'échantillon aléatoire

$$X_1, \ldots, X_n$$

tiré d'une distribution continue F inconnue.

Lorsqu'on veut vérifier si un modèle en particulier F_0 s'ajuste bien aux données, on peut tester

$$\mathcal{H}_0 \, : \, F = F_0, \qquad \mathcal{H}_1 \, : \, F \neq F_0$$

à l'aide d'un test d'adéquation.

Si on rejette $\mathcal{H}_0$, cela veut dire que le modèle F_0 n'est pas adéquat pour les données.

La fonction de répartition empirique

D'abord, notons qu'on peut estimer F à l'aide de la fonction de répartition empirique, définie, pour $x \in \mathbb{R}$ par

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(X_i \leq x).$$

En un point fixé x_0, on a que $F_n(x_0)$ est la proportion échantillonnale d'observations qui sont plus petite ou égale à x_0. On sait donc que, quand $n \to \infty$,

$$F_n(x_0) \xrightarrow{P} \Pr(X_i \leq x_0) = F(x_0)$$

et que

$$\sqrt{n}\{F_n(x) - F(x)\} \rightsquigarrow \mathcal{N}(0, F(x)\{1 - F(x)\})$$

On peut montrer que le processus $\sqrt{n}\{F_n(x) - F(x)\}$ converge en distribution vers un *pont brownien*.

Test de Kolmogorov–Smirnov

Un test d'adéquation classique est basé sur la statistique de Kolmogorov–Smirnov

$$D = \sup_x |F_n(x) - F_0(x)|,$$

c'est-à-dire la distance maximale entre la fonction de répartition empirique et le modèle sous $\mathcal{H}_0$.

Lorsque $n \to \infty$, $F_n(x)$ tend vers la vraie valeur $F(x)$ en chaque point x.

Donc, si $\mathcal{H}_0$ est vraie, on s'attend à voir de petites valeurs de la statistique de Kolmogorov–Smirnov.

La distribution d'échantillonage asymptotique est complexe. Il vaut mieux la déterminer par simulations si on a estimé les paramètres de la loi F_0 à partir des données.

Exemple (suite)

Le test de Kolmogorov–Smirnov est implanté en R. Par exemple, pour tester

$$\mathcal{H}_0 \ : \ F \text{ est la distribution Gamma}(4.4, 1/0.6)$$

dans l'exemple des durées d'accouchement, on trouve :

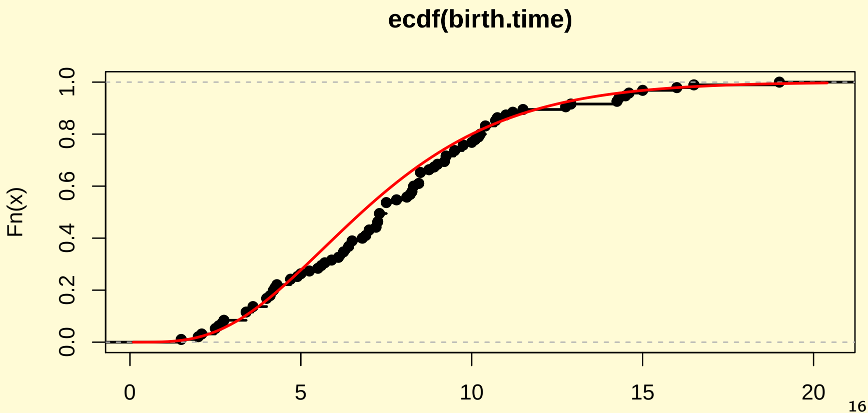
```
ks.test(birth.time,"pgamma", shape = 4.4, rate = 0.6)

##
##  One-sample Kolmogorov-Smirnov test
##
## data:  birth.time
## D = 0.11679, p-value = 0.1497
## alternative hypothesis: two-sided
```

On ne rejette pas $\mathcal{H}_0$ au seuil de 5%, donc on ne peut pas conclure que le modèle Gamma est inadéquat.

Exemple (fin)

On peut voir ici la fonction de répartition empirique comparée avec F_0.

```
plot(ecdf(birth.time), lwd=2)
curve(pgamma(x,shape=4.4,rate=0.6), add=TRUE, col=2, lwd=2)
```



Le test du χ^2 de Pearson

Lorsque les données sont discrètes ou catégorielles, on peut utiliser le test du χ^2 de Pearson pour

- détecter une association entre deux variables,

Le test du χ^2 de Pearson

Lorsque les données sont discrètes ou catégorielles, on peut utiliser le test du χ^2 de Pearson pour

- détecter une association entre deux variables,
- tester l'adéquation d'une distribution discrète.

Design multinomial

1. L'expérience consiste en n essais identiques.

Design multinomial

1. L'expérience consiste en n essais identiques.
2. Le résultat de chaque essai tombe exactement dans une des k catégories (ou cellules).

Design multinomial

1. L'expérience consiste en n essais identiques.
2. Le résultat de chaque essai tombe exactement dans une des k catégories (ou cellules).
3. La probabilité que le résultat d'un essai soit dans la cellule i est π_i, pour $i \in \{1, \ldots, k\}$, et est constante pour toute la durée de l'expérience. On a aussi que

$$\sum_{i=1}^{k} \pi_i = 1.$$

Design multinomial

1. L'expérience consiste en n essais identiques.
2. Le résultat de chaque essai tombe exactement dans une des k catégories (ou cellules).
3. La probabilité que le résultat d'un essai soit dans la cellule i est π_i, pour $i \in \{1, \ldots, k\}$, et est constante pour toute la durée de l'expérience. On a aussi que
$$\sum_{i=1}^{k} \pi_i = 1.$$
4. Les essais sont indépendants.

Design multinomial

1. L'expérience consiste en n essais identiques.
2. Le résultat de chaque essai tombe exactement dans une des k catégories (ou cellules).
3. La probabilité que le résultat d'un essai soit dans la cellule i est π_i, pour $i \in \{1, \ldots, k\}$, et est constante pour toute la durée de l'expérience. On a aussi que

$$\sum_{i=1}^{k} \pi_i = 1.$$

4. Les essais sont indépendants.
5. On s'intéresse à n_i, le nombre d'essais pour lequel le résultat est dans la catégorie i. On a aussi que

$$\sum_{i=1}^{k} n_i = n.$$

Le test du χ^2 de Pearson

On a que

$$X^2 = \sum_{i=1}^{k} \frac{\{n_i - \mathrm{E}(n_i)\}^2}{\mathrm{E}(n_i)} = \sum_{i=1}^{k} \frac{\{n_i - n\pi_i\}^2}{n\pi_i}$$

suit approximativement une distribution χ^2.

Le test du χ^2 de Pearson

On a que

$$X^2 = \sum_{i=1}^{k} \frac{\{n_i - \mathrm{E}(n_i)\}^2}{\mathrm{E}(n_i)} = \sum_{i=1}^{k} \frac{\{n_i - n\pi_i\}^2}{n\pi_i}$$

suit approximativement une distribution χ^2.

Le nombre de degrés de libertés est égal au nombre de classes k moins le nombre de restrictions linéaires placées sur les probabilités.

- Il y a toujours au moins une restriction, soit $\pi_1 + \cdots + \pi_k = 1$.
- Lorsqu'on estime les probabilités $\pi_1, \ldots, \pi_k$, on pose également une (ou des) contrainte(s).

Exemple: test d'adéquation

On modélise le nombre de sinistres en assurance auto pour 100 assurés indépendants et choisis aléatoirement.

On a observé le tableau de fréquences suivant :

```
table(x)

## x
##  0  1  2
## 78 21  1
```

Exemple [suite]

Est-ce que le nombre de sinistres X suit une loi de Poisson ?

On teste l'hypothèse

$$\mathcal{H}_0 : X \sim \text{Poisson}$$

versus

$$\mathcal{H}_1 : X \text{ ne suit pas une Poisson.}$$

On utilise le test du khi-carré de Pearson en supposant 3 cellules: 0, 1, 2 et plus.

Exemple [suite]

Est-ce que le nombre de sinistres X suit une loi de Poisson ?

On teste l'hypothèse

$$\mathcal{H}_0 : X \sim \text{Poisson}$$

versus

$$\mathcal{H}_1 : X \text{ ne suit pas une Poisson.}$$

On utilise le test du khi-carré de Pearson en supposant 3 cellules: 0, 1, 2 et plus.

Premièrement, on ajuste la loi de Poisson avec la méthode du maximum de vraisemblance, ce qui nous donne

$$\hat{\lambda} = \bar{x} = 0.23$$

Exemple [suite]

On calcule les probabilités des cellules selon la loi de Poisson ajustée:

$$\begin{aligned}\hat{\pi}_0 &= \quad e^{-\hat{\lambda}} = 0.7945\\ \hat{\pi}_1 &= \hat{\lambda}e^{-\hat{\lambda}} = 0.1827\\ \hat{\pi}_2 &= 1 - \hat{\pi}_0 - \hat{\pi}_1 = 0.0227.\end{aligned}$$

On trouve donc que les fréquences espérées selon le modèle sont

$$\begin{aligned}\widehat{E(n_0)} &= n\hat{\pi}_0 = 79.45\\ \widehat{E(n_1)} &= n\hat{\pi}_1 = 18.27\\ \widehat{E(n_2)} &= n\hat{\pi}_2 = 2.27.\end{aligned}$$

Exemple [fin]

La statistique du test est donc

$$\begin{aligned} X^2 &= \sum_{i=0}^{2} \frac{\{n_i - \widehat{E(n_i)}\}^2}{\widehat{E(n_i)}} \\ &= \frac{(78 - 79.45)^2}{79.45} + \frac{(21 - 18.27)^2}{18.27} + \frac{(1 - 2.27)^2}{2.27} \\ &= 1.145 \end{aligned}$$

Sous $\mathcal{H}_0$, X^2 suit une χ^2 avec $k - 2 = 1$ degré de liberté. Puisque $\chi^2_{1,5\ \%} = 3.84$, on ne rejette pas l'hypothèse nulle que la distribution est Poisson.

Tableau de contingence

Plusieurs problèmes statistiques peuvent être résolus à l'aide de tableaux de contingence. Dans un tel tableau, on classifie des sujets selon (au moins) deux variables catégorielles avec chacune (au moins) deux niveaux.

Tableau de contingence

Plusieurs problèmes statistiques peuvent être résolus à l'aide de tableaux de contingence. Dans un tel tableau, on classifie des sujets selon (au moins) deux variables catégorielles avec chacune (au moins) deux niveaux.

Par exemple, on peut compter le nombre

- de personnes selon la couleur de leurs yeux et leur cheveux,
- de personnes selon leur sexe et leur opinion politique,
- de patients selon la dose de médicament et le taux de rémission,
- etc.

Tableau de contingence

Plusieurs problèmes statistiques peuvent être résolus à l'aide de tableaux de contingence. Dans un tel tableau, on classifie des sujets selon (au moins) deux variables catégorielles avec chacune (au moins) deux niveaux.

Par exemple, on peut compter le nombre

- de personnes selon la couleur de leurs yeux et leur cheveux,
- de personnes selon leur sexe et leur opinion politique,
- de patients selon la dose de médicament et le taux de rémission,
- etc.

Dans ces exemples, on veut investiguer s'il existe une dépendance entre les deux critères de classification.

Association entre la couleur des cheveux et des yeux

Y a-t-il un lien entre la couleur des yeux et la couleur des cheveux?
On classifie 592 étudiants de statistique à l'Université de Delaware[1] selon la couleur de leurs cheveux et de leurs yeux dans le tableau de contingence suivant :

```
##          Hair
## Eye       Black Brown Red Blond Sum
##   Brown      68   119  26     7 220
##   Blue       20    84  17    94 215
##   Hazel      15    54  14    10  93
##   Green       5    29  14    16  64
##   Sum       108   286  71   127 592
```

[1] Snee, R. D. (1974). Graphical display of two-way contingency tables. *The American Statistician*, **28**, 9–12.

Association entre la couleur des cheveux et des yeux

On veut tester

$\mathcal{H}_0$: la couleur des yeux ne dépend pas de la couleur des cheveux

Association entre la couleur des cheveux et des yeux

On veut tester

$\mathcal{H}_0$: la couleur des yeux ne dépend pas de la couleur des cheveux

- Soient p_{noir}, p_{brun}, p_{roux}, p_{blond} les probabilités d'avoir les cheveux noirs, bruns, roux ou blonds, respectivement (les probabilités des colonnes). On a

$$p_{noir} + p_{brun} + p_{roux} + p_{blond} = 1.$$

Association entre la couleur des cheveux et des yeux

On veut tester

$\mathcal{H}_0$: la couleur des yeux ne dépend pas de la couleur des cheveux

- Soient p_{noir}, p_{brun}, p_{roux}, p_{blond} les probabilités d'avoir les cheveux noirs, bruns, roux ou blonds, respectivement (les probabilités des colonnes). On a

$$p_{noir} + p_{brun} + p_{roux} + p_{blond} = 1.$$

- Soient q_{brun}, q_{bleu}, $q_{noisette}$, q_{vert} les probabilités d'avoir les yeux bruns, bleus, noisette ou verts, respectivement (les probabilités des lignes). On a

$$q_{brun} + q_{bleu} + q_{noisette} + q_{vert} = 1.$$

Probabilités des cellules sous $\mathcal{H}_0$

Si $\mathcal{H}_0$ est vraie, alors la probabilité de la cellule i, j est simplement le produit $q_i p_j$.

Probabilités des cellules sous $\mathcal{H}_0$

Si $\mathcal{H}_0$ est vraie, alors la probabilité de la cellule i, j est simplement le produit $q_i p_j$.
Par exemple, la probabilité d'avoir les yeux bruns et les cheveux noirs serait

$$q_{brun} \times p_{noir}.$$

Probabilités des cellules sous $\mathcal{H}_0$

Si $\mathcal{H}_0$ est vraie, alors la probabilité de la cellule i, j est simplement le produit $q_i p_j$.
Par exemple, la probabilité d'avoir les yeux bruns et les cheveux noirs serait

$$q_{brun} \times p_{noir}.$$

On peut estimer les probabilités des lignes et des colonnes à l'aide des proportions empiriques. Si r_i est le nombre d'observations dans la ligne i et c_j est le nombre d'observations dans la colonne j, on a

$$\hat{q}_i = r_i/n \quad \text{et} \quad \hat{p}_j = c_j/n.$$

Alors, le nombre espéré dans la case i, j est

$$\widehat{\mathrm{E}(n_{ij})} = n\hat{q}_i\hat{p}_j = \frac{r_i c_j}{n}.$$

Test d'indépendance du χ^2

La statistique du test du χ^2 de Pearson est

$$X^2 = \sum_{j=1}^{4} \sum_{i=1}^{4} \frac{\{n_{ij} - \widehat{\mathrm{E}(n_{ij})}\}^2}{\widehat{\mathrm{E}(n_{ij})}}.$$

Sous $\mathcal{H}_0$ dans un tableau de dimension $r \times c$, cette statistique sera approximativement χ^2 avec degrés de liberté

$$rc - 1 - (r-1) - (c-1) = (r-1)(c-1).$$

Exemple : Comptes espérés

Dans l'exemple des couleurs de cheveux et de yeux, on trouve donc les comptes espérés ci-dessous

```
##        [,1]   [,2]  [,3]  [,4]
## [1,] 40.14 106.28 26.39 47.20
## [2,] 39.22 103.87 25.79 46.12
## [3,] 16.97  44.93 11.15 19.95
## [4,] 11.68  30.92  7.68 13.73
```

Exemple : Conclusion

La valeur observée de la statistique et le seuil observé, calculé avec 3×3 degrés de liberté sont donc

```
(Xstat <- sum((tableau-Ei)^2/Ei))

## [1] 138.2925

pchisq(Xstat,df=9,lower.tail=FALSE)

## [1] 2.322332e-25
```

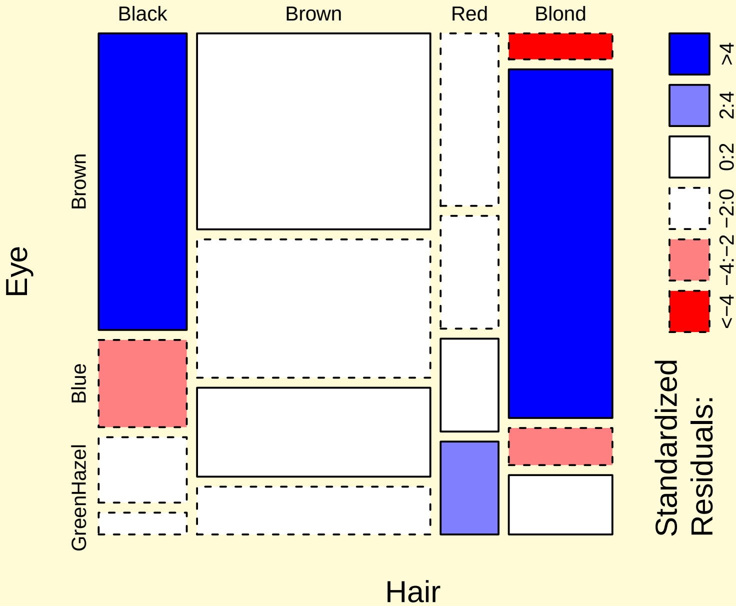
On rejette l'hypothèse d'indépendance entre la couleur des yeux et des cheveux. La couleur des yeux dépend de celle des cheveux.

Diagramme en mosaïque

La fonction `mosaicplot` permet de visualiser ce test d'indépendance.

```
mosaicplot(~Hair+Eye, data=HairEyeColor, shade=TRUE,
           type="pearson")
```

Exemple : Visualisation



Exemple : Visualisation

Dans le graphique précédent, l'aire de la case est proportionnelle à la fréquence dans la case.

Exemple : Visualisation

Dans le graphique précédent, l'aire de la case est proportionnelle à la fréquence dans la case.

Les cases bleues ou avec une ligne pleine représentent les cases pour lesquelles $n_{ij} - \widehat{\mathrm{E}(n_{ij})} > 0$, donc pour lesquelles on a un compte plus élevé que prévu par le modèle d'indépendance.

Exemple : Visualisation

Dans le graphique précédent, l'aire de la case est proportionnelle à la fréquence dans la case.

Les cases bleues ou avec une ligne pleine représentent les cases pour lesquelles $n_{ij} - \widehat{\mathrm{E}(n_{ij})} > 0$, donc pour lesquelles on a un compte plus élevé que prévu par le modèle d'indépendance.

Les cases rouges ou avec une ligne pointillée sont celles pour lesquelles $n_{ij} - \widehat{\mathrm{E}(n_{ij})} < 0$, donc le compte est moins élevé que prévu par le modèle d'indépendance.