Theory of statistics 2

Department of Statistics and Operations Research

March 23, 2020

Composite Hypothesis

The general shape of composite hypothesis is: $H_0 : \theta \in \Omega_0$ vs $H_1 : \theta \in \Omega_1$. For any test γ, we defined its power function $\pi_\gamma(\theta) = \mathbb{P}(rejH_0|\theta)$ and its size $\alpha = \sup_{\theta \in \Omega_0} (\pi_\gamma(\theta))$. The probability of the error type 2 is the function $\beta(\theta) = 1 - \pi_\gamma(\theta)_{\theta \in \Omega_1}$. In specific, three cases are distributed as follow:

- $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$.

- $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$.

- $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$.

Three tests will be generated

1. The Uniformly Most Powerful Test γ_{UMP} of size α_{UMP}.
2. The Generalized Likelihood Ratio Test γ_{GLR} of size α_{GLR}.
3. The test γ_{CI} of size α based on $100(1-\alpha)\%$ C.I for θ_0.

More precisely, test γ_{UMP} will be used for the cases $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ and $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$, whereas γ_{GLR} and γ_{CI} will be used for the case $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$.

Uniformly Most Powerful Test

Definition

The Uniformly Most Powerful Test γ_{UMP} of size α_{UMP} for $H_0 : \theta \in \Omega_0$ vs $H_1 : \theta \in \Omega_1$ satisfies: for all other test γ with size α, we have

$$\alpha \leq \alpha_{UMP}.$$

Consequently, $\pi_{\gamma_{UMP}}(\theta)_{\theta \in \Omega_1} \geq \pi_{\gamma}(\theta)_{\theta \in \Omega_1}$ ($\beta_{UMP} \leq \beta$, for all $\theta \in \Omega_1$).

Theorem

If $f(x;\theta)$ belongs to the class of exponential families, then the test γ_{UMP} for the cases $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$ and $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$ is similar to the test γ_{MP} for $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$ rejects H_0 is reduced as follows:

	$H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$	$H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$
$c(\theta) \nearrow$	$\sum d(x_i) > k$	$\sum d(x_i) < k$
$c(\theta) \searrow$	$\sum d(x_i) < k$	$\sum d(x_i) > k$
	Alike	Inverse

k solves the equation

$$\alpha_{UMP} = \mathbf{P}\left(Reject \quad H_0 | \theta_0\right).$$

Example 1: The normal distribution with known variance

Let $X_1, \ldots, X_n$ be a sample of size $n = 25$ that drawn from a normal distribution $N(\theta, 16)$. Our aim is to test the γ_{UMP} test with size $\alpha = 0.05$ for $H_0 : \theta \leq 3$ vs $H_1 : \theta > 3$. Since $f(x;\theta) \in$Exp. Fam with $C(\theta) = \frac{\theta}{\sigma^2} \nearrow$, then γ_{UMP} rejects H_0 if $\sum d(x_i) = \sum x_i > k$ or $\overline{X} > C$. It follows that

$$0.05 = \mathbb{P}\left(\overline{X} > C | \theta = 3\right) = \mathbb{P}\left(Z > \frac{C-3}{4/5}\right) \Rightarrow \frac{C-3}{4/5} = 1.645.$$

This implies that

$$C = 4.316 \Rightarrow k = 25C.$$

- Homework

Find γ_{UMP} test with size $\alpha = 0.05$ for $H_0 : \theta \geq 3$ vs $H_1 : \theta < 3$.

Example 2: The exponential distribution

Let $X_1, \ldots, X_n$ be a sample of size $n = 10$ that drawn from an exponential distribution $\exp(\theta)$. Our aim is to test the γ_{UMP} test with size $\alpha = 0.05$ for $H_0 : \theta \leq 2$ vs $H_1 : \theta > 2$. Since $f(x;\theta) \in$Exp. Fam with $C(\theta) = -\theta \searrow$, then γ_{UMP} rejects H_0 if $S = \sum x_i < k$, where k is found by solving the equation

$$0.05 = \mathbb{P}\left(S < k|\theta = 2\right) = \mathbb{P}\left(U < 4k\right),$$

where $U = 2\theta S \sim \mathcal{X}^2_{2n}$. Then

$$4k = 10.851 \Rightarrow k = 2.713.$$

- Homework

Find γ_{UMP} test with size $\alpha = 0.05$ for $H_0 : \theta \geq 2$ vs $H_1 : \theta < 2$.

Example 3: The normal distribution with known mean

Let $X_1, \ldots, X_n$ be a sample of size $n = 10$ that drawn from a normal distribution $N(\theta, \sigma^2)$. Our aim is to test the γ_{UMP} test with size $\alpha = 0.05$ for $H_0 : \sigma \leq 2$ vs $H_1 : \sigma > 2$. Since $f(x; \theta) \in$Exp. Fam with $C(\sigma) = -\frac{1}{2}\frac{1}{\sigma^2} \nearrow$, then γ_{UMP} rejects H_0 if $\sum d(x_i) = \sum (x_i - \theta)^2 > k$. It follows that

$$0.05 = \mathbb{P}\left(\sum (x_i - \theta)^2 > k | \sigma = 2\right) = \mathbb{P}\left(\sum \left(\frac{x_i - \theta}{2}\right)^2 > \frac{k}{4}\right).$$

This implies that

$$\frac{k}{4} = 18.31 \Rightarrow k = 73.24.$$

- Homework

Find γ_{UMP} test with size $\alpha = 0.05$ for $H_0 : \sigma \geq 2$ vs $H_1 : \sigma < 2$.

The test GLR γ_{GLR} is used for case: $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$.

Definition

The test GLR γ_{GLR} of size α for the composite hypothesis:

$$H_0 : \theta \in \Omega_0 \text{ vs } H_1 : \theta \in \Omega_1.$$

is found by the steps:

1. Take the Generalized Likelihood Ratio (GLR):

$$\lambda = \frac{\sup_{\theta \in \Omega_0} \ell(\underline{X}, \theta)}{\sup_{\theta \in \Omega} \ell(\underline{X}, \theta)}, \quad \Omega = \Omega_0 \cup \Omega_1.$$

2. Reject H_0 if $\lambda < k$.
3. Find k by solving the implicit equation: $\alpha = \sup_{\theta \in \Omega_0} \mathbb{P}(\lambda < k | \theta)$.

Indeed, maximizing the nominator and the denominator of λ depends on the MLE of θ in Ω_0 and in Ω, respectively. In the case $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, γ_{GLR} rejects H_0 if

$$\lambda = \frac{\ell(\underline{X}, \theta_0)}{\ell(\underline{X}, \widehat{\theta}_{MLE})} < k, \text{ where } k \text{ solves } \alpha = \sup_{\theta \in \Omega_0} \mathbb{P}(\lambda < k | \theta).$$

Example 1

Find γ_{GLR} and α for $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ in the case $f(x;\theta) = N(\theta, \sigma^2)$, where σ^2 is unknown.
The estimators $\widehat{\theta}_{MLE} = \overline{X}$ and $\widehat{\sigma^2}_{MLE} = \frac{1}{n}\sum(X_i - \overline{X})^2$. γ_{GLR} rejects H_0 if $\lambda = \dfrac{\ell(\underline{X}, \theta_0)}{\ell(\underline{X}, \overline{X})} < k$. In fact,

$$\begin{aligned}\lambda &= \frac{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2}\frac{\sum(X_i-\theta_0)^2}{\sigma^2}\right)}{\left(\frac{1}{\sqrt{2\pi}\widehat{\sigma}_{MLE}}\right)^n \exp\left(-\frac{1}{2}\frac{\sum(X_i-\widehat{\theta}_{MLE})^2}{\widehat{\sigma}^2_{MLE}}\right)} \\ &= \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{n}\sum(X_i-\theta_0)^2\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\frac{\sum(X_i-\theta_0)^2}{\frac{1}{n}\sum(X_i-\theta_0)^2}\right)}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{n}\sum(X_i-\overline{X})^2\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\frac{\sum(X_i-\overline{X})^2}{\frac{1}{n}\sum(X_i-\overline{X})^2}\right)}\end{aligned}$$

It follows that

$$\lambda = \frac{\left(\frac{1}{n}\sum(X_i-\theta_0)^2\right)^{-\frac{n}{2}}}{\left(\frac{1}{n}\sum(X_i-\overline{X})^2\right)^{-\frac{n}{2}}} = \left(\frac{\sum(X_i-\overline{X})^2}{\sum(X_i-\theta_0)^2}\right)^{\frac{n}{2}} < k.$$

Thus

$$\frac{\sum(X_i-\overline{X})^2}{\sum(X_i-\theta_0)^2} = \frac{\sum(X_i-\overline{X})^2}{\sum(X_i-\overline{X})^2 + n(\overline{X}-\theta_0)^2} < k^{\frac{2}{n}} = k_1.$$

This implies that

$$\frac{n(\overline{X}-\theta_0)^2}{\sum(X_i-\overline{X})^2} > \frac{1}{k_1} - 1 = k_2 \Leftrightarrow \left(\frac{\overline{X}-\theta_0}{S/\sqrt{n}}\right)^2 > c^2.$$

Then, γ_{GLR} rejects H_0 if

$$\begin{aligned}\alpha &= \mathbb{P}(\lambda < k|\theta_0) = \mathbb{P}\left(\left(\frac{\overline{X}-\theta_0}{S/\sqrt{n}}\right)^2 > c^2\right) \\ &= \mathbb{P}\left(\frac{\overline{X}-\theta_0}{S/\sqrt{n}} > c\right) + \mathbb{P}\left(\frac{\overline{X}-\theta_0}{S/\sqrt{n}} < -c\right).\end{aligned}$$

For $n < 30$, $c = t_{\frac{\alpha}{2},n-1}$ and for $n \geq 30$, we have $c = z_{\frac{\alpha}{2}}$.
We accept H_0 if

$$\begin{cases} n < 30 \Rightarrow \frac{\overline{X}-\theta_0}{S/\sqrt{n}} \in \left(-t_{\frac{\alpha}{2},n-1}, t_{\frac{\alpha}{2},n-1}\right) \\ n \geq 30 \Rightarrow \frac{\overline{X}-\theta_0}{S/\sqrt{n}} \in \left(-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}}\right) \end{cases}$$

Example 2

Find γ_{GLR} and α for $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ in the case $f(x; \theta) = \exp(\theta)$.

The estimators $\widehat{\theta}_{MLE} = \frac{1}{\overline{X}}$. γ_{GLR} rejects H_0 if $\lambda = \dfrac{\ell(\underline{X}, \theta_0)}{\ell(\underline{X}, \overline{X})} < k$.

In fact,

$$\lambda = \frac{\theta_0^n \exp(-\theta_0 \sum x_i)}{\left(\frac{1}{\overline{X}}\right)^n \exp\left(-\frac{\sum x_i}{\overline{X}}\right)} = (\theta_0 S)^n \left(\frac{e}{n}\right)^n \exp\left(-\theta_0 S\right) < k,$$

where $S = \sum x_i$. Since $U = 2\theta_0 S \sim \mathcal{X}^2_{2n}$, then

$$(2\theta_0 S)^n \exp\left(-\frac{2\theta_0 S}{2}\right) < k \left(\frac{2n}{e}\right)^n = c.$$

Thus

$$f(u) = u^n e^{-u/2} < c, \text{ where } c \text{ sloves } \mathbb{P}(f(u) < c) = \alpha.$$

The Confidence Interval Test (C.I.) γ_{CI} based on $100(1-\alpha)$C.I. is suitable to test the case:

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta \neq \theta_0.$$

Simply, if $100(1-\alpha)$C.I. for θ_0 is $(T_1(\underline{X}), T_2(\underline{X}))$, then γ_{CI} accepts H_0 if:

$$\theta_0 \in (T_1(\underline{X}), T_2(\underline{X}))\,.$$

Note that the tests γ_{GLR} and γ_{CI} can coincide.

Example 1

Find γ_{CI} and α for $H_0 : \theta = \theta_0$ vs $H_0 : \theta \neq \theta_0$ in the case $f(x;\theta) = N(\theta, \sigma^2)$, where σ^2 is unknown.
$100(1-\alpha)$C.I. is

$$(T_1(\underline{X}), T_2(\underline{X})) = \begin{cases} n < 30 \Longrightarrow \left(\overline{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}\right) \\ n \geq 30 \Longrightarrow \left(\overline{X} \pm z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right) \end{cases}$$

Then, we accept H_0 if

$$n < 30 \Rightarrow \theta_0 \in \left(\overline{X} \pm t_{1-\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}\right) \Leftrightarrow \frac{\overline{X} - \theta_0}{S/\sqrt{n}} \in \left(-t_{1-\frac{\alpha}{2}, n-1}, t_{1-\frac{\alpha}{2}, n-1}\right)$$

$$n \geq 30 \Rightarrow \theta_0 \in \left(\overline{X} \pm z_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right) \Leftrightarrow \frac{\overline{X} - \theta_0}{S/\sqrt{n}} \in \left(-z_{1-\frac{\alpha}{2}}, z_{1-\frac{\alpha}{2}}\right).$$

Note that γ_{CI} and γ_{GLR} are similar.

Example 2

Find γ_{CI} and α for $H_0 : \sigma^2 = \sigma_0^2$ vs $H_0 : \sigma^2 \neq \sigma_0^2$ in the case $f(x; \sigma^2) = N(\mu, \sigma^2)$, where μ is known.
$100(1-\alpha)$C.I. is

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\mathcal{X}^2_{n,\frac{\alpha}{2}}}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\mathcal{X}^2_{n,1-\frac{\alpha}{2}}} \right).$$

Then, we accept H_0 if

$$\sigma_0^2 \in \left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\mathcal{X}^2_{n,\frac{\alpha}{2}}}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\mathcal{X}^2_{n,1-\frac{\alpha}{2}}} \right)$$

or

$$\sum_{i=1}^n (X_i - \mu)^2 \in \left(\sigma_0^2 \mathcal{X}^2_{n,1-\frac{\alpha}{2}}, \sigma_0^2 \mathcal{X}^2_{n,\frac{\alpha}{2}} \right).$$

If n is large enough, the approximated C.I. is equal to

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(\widehat{\sigma}^2_{MLE} \pm \widehat{\sigma}^2_{MLE} z_{1-\frac{\alpha}{2}} \sqrt{\frac{2}{n}}\right),$$

where $\widehat{\sigma}^2_{MLE} = \frac{1}{n}\sum(x_i - \mu)^2$. Then, we accept H_0 if

$$\sigma_0^2 \in \left(\widehat{\sigma}^2_{MLE} \pm \widehat{\sigma}^2_{MLE} z_{1-\frac{\alpha}{2}} \sqrt{\frac{2}{n}}\right)$$

or

$$\sum_{i=1}^{n}(X_i - \mu)^2 \in \left(\frac{n\sigma_0^2}{1 + z_{1-\frac{\alpha}{2}}\sqrt{\frac{2}{n}}}, \frac{n\sigma_0^2}{1 - z_{1-\frac{\alpha}{2}}\sqrt{\frac{2}{n}}}\right).$$

Example 3

Find γ_{GLR} and α for $H_0 : \theta = \theta_0$ vs $H_0 : \theta \neq \theta_0$ in the case $f(x;\theta) = \exp(-(x-\theta)), \quad x > \theta$. $100(1-\alpha)$C.I. is

$$(T_1(\underline{X}), T_2(\underline{X})) = \left(S + \frac{\log(\alpha)}{n}, S\right), \text{ where } S = \sum x_i.$$

Then, we accept H_0 if

$$\theta_0 \in \left(S + \frac{\log(\alpha)}{n}, S\right) \text{ or } S \in \left(\theta_0, \theta_0 - \frac{\log(\alpha)}{n}\right).$$

Homework

- Find γ_{CI} and α for $H_0 : \theta = \theta_0$ vs $H_0 : \theta \neq \theta_0$ in the case where $f(x;\theta) = \dfrac{1}{\theta}, \quad 0 < x < \theta.$

- Find γ_{CI} and α for $H_0 : \theta = \theta_0$ vs $H_0 : \theta \neq \theta_0$ in the case where $f(x;\theta) = \theta \exp(-\theta x), \quad 0 < x.$

Thank you