Introduction to Biostatistical Theory - Inference

Robert Parker

Elements of a Hypothesis Test

A hypothesis test is a statement about a population parameter.

Given data $(X_1, X_2, \ldots, X_n) \sim f(\boldsymbol{x}|\theta)$, we wish to make a claim about the parameter $\theta \in \Theta$. We test this by comparing two hypotheses about the true value of the parameter θ:

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

where $\Theta_0 \cap \Theta_1 = \emptyset$. (Note: $\Theta_0 \cup \Theta_1$ need not equal Θ.)

- H_0 is called the **null hyphothesis**.
- H_1 (or H_A) is called the **alternative hypothesis**

Elements of a Hypothesis Test - simple vs Composite

1. If you are interested in if a coin is fair, we could toss it 1000 times and get 600 heads (say). Let $p = P(\text{head})$ and the parameter space be $\Theta = [0,1]$. We would then want to test

$$H_0 : p = \frac{1}{2} \quad \text{vs} \quad H_A : p \neq \frac{1}{2} \; (p \in [0, 1/2) \cup (1/2, 1]).$$

   (The alternative is a union of two intervals and is called a **composite hypothesis**.)

2. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$, θ unknown. We could test

$$H_0 : \theta = 0 \;\; \text{vs} \; H_1 : \theta = 1.$$

   Both the null and alternative hypotheses are examples of a **simple hypothesis** (testing against a single value).

Elements of a Hypothesis Test - Choosing the Hypotheses

Let $\theta =$ the proportion of defective items, θ_0 be the maximum acceptable proportion of defective items. In this case, we could test

$$H_0 : \theta \geq \theta_0 \quad \text{vs} \quad H_A : \theta < \theta_0.$$

Here, the alternative hypothesis is what we are most interested in.

In statistics, people frequently specify:

- Null hypothesis: "no change", "no difference", or "no effect"
- Alternative hypothesis: "some effect", "some differences", or "some change"

In general, the alternative hypothesis is the statement that exhibits a difference from normal or the condition that is of interest.

Elements of a Hypothesis Test - General Idea

Definition

A hypothesis testing procedure or hypothesis test is a rule that specifies

1. for which sample values the decision is made to accept H_0 as true
2. for which sample values H_0 is rejected and H_1 is accepted as true.

Suppose wish to test

$$H_0 : \theta = \Theta_0 \;\text{ vs }\; H_1 : \theta \in \Theta_1.$$

Assume H_0 is true. We can make our conclusion based off of whether or not the observed data is "likely" or "not likely" under this assumption. If the observed data is "likely" under H_0, then we don't reject H_0 (there is not enough evidence to reject it). If the observed data is "not likely", then we reject H_0.

Elements of a Hypothesis Test - General Idea

Example: Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$, $\theta \in \mathbb{R}$ unknown. We wish to test

$$H_0 : \theta = 0 \ \text{ vs } \ H_1 : \theta \neq 0.$$

Supposed we observed data such that $\bar{X} = 1.1$ with a sample size of $n = 9$. Assume $H_0 : \theta = 0$ is true. Then

$$\bar{X} \sim N(0, 1/9).$$

What values of $\bar{X}$ are likely under H_0?

Is our observed value of $\bar{X} = 1.1$ likely under H_0?

$$\begin{aligned} P(|\bar{X}| > 1.1) &= 2P(\bar{X} > 1.1) \\ &= 2P\left(\frac{\bar{X} - 0}{1/3} > \frac{1.1 - 0}{1/3}\right) \\ &= 2P(Z > 3.3) = 0.00096 = 0.096\%, \end{aligned}$$

where $Z \sim N(0, 1)$.

Elements of a Hypothesis Test - Critical Region

Definition

Denote the sample space by $\mathcal{D}$. A test of H_0 vs H_1 is based on a subset $\mathcal{R}$ of $\mathcal{D}$. The set $\mathcal{R}$ is called the **critical region** (or (**rejection region**) and its decision rule is reject H_0 if $\boldsymbol{X} \in \mathcal{R}$ and retain H_0 if $\boldsymbol{X} \in \mathcal{R}^c$.

- In the previous example, $\mathcal{D} = \mathbb{R}$ and $\mathcal{R} = \{\bar{X} : |\bar{X}| > 2/3\}$ and $\mathcal{R}^c = \{\bar{X} : |\bar{X}| \leq 2/3\}$.
- Typically, the rejection region will be stated in terms of a test statistic $T(\boldsymbol{X})$ that contains the information about θ and the null is rejected if $T(\boldsymbol{X}) \in \mathcal{R}$.

Elements of a Hypothesis Test - Errors

This procedure can lead to two types of errors:

Decision	Truth: H_0	Truth: H_1
Accept H_0	Correct	Type II error
Reject H_0	Type I error	Correct

- $\alpha = P(\text{Type I Error}) = P(\text{reject } H_0 | H_0 \text{ is true})$
- $\beta = P(\text{Type II Error}) = P(\text{accept } H_0 | H_1 \text{ is true})$
- $1 - \beta = P(\text{reject } H_0 | H_1 \text{ is true})$ (statistical power)

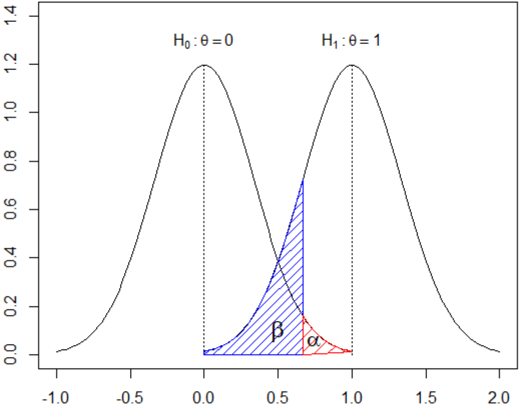
Elements of a Hypothesis Test - Errors

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$, $\theta \in \mathbb{R}$ unknown. We want to test the simple hypotheses

$$H_0 : \theta = 0 \ \text{ vs } \ H_1 : \theta = \theta_1 \ (\theta_1 \neq 0).$$

Assume $n = 9$ as we had before. $\bar{X}$ is sufficient for θ, so let's base our test on this statistic with rejection region $\mathcal{R} = \{\boldsymbol{X} : |\bar{X}| > 2/3\}$.

Elements of a Hypothesis Test - Errors



Elements of a Hypothesis Test

Definition

We say a critical region $\mathcal{R}$ is of **size** α if

$$\sup_{\theta \in \Theta_0} P_\theta(\boldsymbol{X} \in \mathcal{R}) = \alpha.$$

We say a critical region $\mathcal{R}$ is of **level** α if

$$\sup_{\theta \in \Theta_0} P_\theta(\boldsymbol{X} \in \mathcal{R}) \leq \alpha.$$

Definition

The power function of a hypothesis test with rejection region $\mathcal{R}$ is the function of θ defined by

$$\beta(\theta) = P_\theta(\boldsymbol{X} \in \mathcal{R}).$$

- When $\theta \in \Theta_0$, then $\beta(\theta) = \alpha = P(\text{Type I Error})$.
- When $\theta \in \Theta_1$, then $\beta(\theta) = 1 - P(\text{Type II Error})$.

Elements of a Hypothesis Test - Errors

Suppose that the manufacturer of a new medication wants to test the null hypothesis $\theta = 0.9$ against the alternative $\theta = 0.6$. His test statistic is $X = \#$ of successses (recoveries) in 20 trials and he will accept H_0 if $X > 14$ otherwise he will reject H_0. Find the type I and type II error rates.

Elements of a Hypothesis Test - Errors

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$.

$$H_0 : \mu = \mu_0 \;\text{ vs }\; H_1 : \mu = \mu_1 \; (\mu_1 > \mu_0).$$

Find the value k, such that $\bar{X} > k$ provides a critical region of size $\alpha = 0.05$ for a sample of size n and find the power function $\beta(\mu)$.

Elements of a Hypothesis Test - Errors

Generalized Likelihood Ratio Test (GLRT)

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f(x|\theta)$. Recall that the likelihood function is given by

$$L(\theta|\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i|\theta).$$

The likelihood ratio test statistic for testing

$$H_0 : \theta \in \Theta_0 \ \text{ vs } \ H_1 : \theta \in \Theta_1$$

is

$$\lambda(\boldsymbol{X}) = \frac{\sup_{\theta\in\Theta_0} L(\theta|\boldsymbol{X})}{\sup_{\theta\in\Theta} L(\theta|\boldsymbol{X})} = \frac{L(\widehat{\theta}_0|\boldsymbol{X})}{L(\widehat{\theta}|\boldsymbol{X})},$$

where $\Theta = \Theta_0 \cup \Theta_1$, $\widehat{\theta}_0$ is the MLE of $\theta \in \Theta_0$ and $\widehat{\theta}$ is the MLE of $\theta \in \Theta$.

The LRT is any test that has a rejection region of the form $\{\boldsymbol{X} : \lambda(\boldsymbol{X}) \leq C\}$, where C is any number satisfying $0 \leq C \leq 1$.

LRT - Normal Mean With Known Variance

Let $X_1, X_2, \ldots, X_n \stackrel{iid}{\sim} N(\theta, 1)$. Find a size α LRT of

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

LRT - Normal Mean With Known Variance

LRT - Normal Mean With Known Variance

LRT - Normal Mean With Known Variance

LRT - Truncated Exponential

Let $X_1, X_2, \ldots, X_n$ be an iid sample from a truncated exponential distribution with pdf

$$f(x|\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta \\ 0, & \text{otherwise.} \end{cases}$$

Construct a level α LRT test of

$$H_0 : \theta \leq \theta_0 \;\; \text{vs} \;\; H_1 : \theta > \theta_0.$$

LRT - Truncated Exponential

LRT - Truncated Exponential

LRT - Uniform

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$, and suppose we want to test

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

Find the LRT.

LRT and Sufficient Statistics

Let $T(\boldsymbol{X})$ be a sufficient statistic for θ with pdf $g(t|\theta)$ and let the likelihood based on T be given by

$$L^\star(\theta|t) = g(t|\theta)$$

instead of using the sample $\boldsymbol{X}$ and its likelihood $L(\theta|\boldsymbol{x})$. Let $\lambda^\star(t)$ be the LRT statistic based on T. We have $\lambda^\star(T(\boldsymbol{x})) = \lambda(\boldsymbol{x})$ for all samples $\boldsymbol{x}$. The simplified expression for $\lambda(\boldsymbol{X})$ should depend on $\boldsymbol{x}$ only through $T(\boldsymbol{x})$ if $T(\boldsymbol{x})$ is a sufficient statistic for θ.

LRT - Normal Mean With Known Variance

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$. Find a size α LRT of

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

LRT - Truncated Exponential

Let $X_1, X_2, \ldots, X_n$ be an iid sample from a truncated exponential distribution with pdf

$$f(x|\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta \\ 0, & \text{otherwise.} \end{cases}$$

Construct a level α LRT test of

$$H_0 : \theta \leq \theta_0 \;\text{ vs }\; H_1 : \theta > \theta_0.$$

LRT - Truncated Exponential

LRT - Uniform

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$, and suppose we want to test

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

Find the LRT.

Uniformly Most Powerful Tests (UMP)

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f(x|\theta)$ and suppose that we want to test

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

using a test with rejection region $\mathcal{R}$, i.e. reject if $\boldsymbol{X} \in \mathcal{R}$.

Each test is associated with a rejection region $\mathcal{R}$. We may define a test function

$$\phi(\boldsymbol{X}) = \begin{cases} 1, & \boldsymbol{X} \in \mathcal{R} \\ 0, & \boldsymbol{X} \notin \mathcal{R} \end{cases}$$

This is a 1-1 mapping between ϕ and $\mathcal{R}$. The power function is $\beta_\phi(\theta) = P_\theta(\phi(\boldsymbol{X}) = 1) = E_\theta[\phi(\boldsymbol{X})]$. In general, a statistical hypothesis test is a test function $\phi : \mathcal{X} \to [0, 1]$ such that if $\boldsymbol{X} = \boldsymbol{x}$ is observed you reject the null hypothesis with probability $\phi(\boldsymbol{x})$ and accept the null hypothesis with probability $1 - \phi(\boldsymbol{x})$, and the power function is defined by $\beta(\theta) = E_\theta[\phi(\boldsymbol{X})]$

Uniformly Most Powerful Tests (UMP)

A test ϕ is a UMP (uniformly most powerful) test with level α of

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

if $E_\theta[\phi(\boldsymbol{X})] \leq \alpha$ for all $\theta \in \Theta_0$ and $\beta_\phi(\theta) \geq \beta_{\phi'}(\theta)$ for all $\theta \in \Theta_1$ where $\phi' \in \mathcal{C}$ and $\mathcal{C} = \{\phi : E_\theta[\phi(\boldsymbol{X})] \leq \alpha \text{ for all } \theta \in \Theta_0\}$.

Neyman-Pearson Lemma

Theorem (Neyman-Pearson Lemma)

Consider testing

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$$

where the pdf or pmf corresponding to θ_i *is given by* $f(x|\theta_i), i = 0, 1$. *Define a test* $\phi(\boldsymbol{X})$ *satisfying* $E_{\theta_0}[\phi(\boldsymbol{X})] = \alpha$ (1) *and for some* $k > 0$ *and* $0 \leq c < 1$

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } f(\boldsymbol{x}|\theta_1) > kf(\boldsymbol{x}|\theta_0) \\ 0, & \text{if } f(\boldsymbol{x}|\theta_1) < kf(\boldsymbol{x}|\theta_0) \end{cases} = \begin{cases} 1, & \text{if } \dfrac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} > k \\ 0, & \text{if } \dfrac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} < k \end{cases} \quad (2)$$

then

(a) $\phi(\boldsymbol{X})$ *is the UMP level* α *test. (Here, rather than UMP, we say most powerful since* $\Theta_1 = \{\theta_1\}$.)

(b) *if* $\phi(\boldsymbol{X})$ *is UMP level* α *test then there exists a constant* k *such that* $\phi(\boldsymbol{X})$ *satisfies* (1) *and* (2).

Neyman-Pearson Lemma - Proof

Neyman-Pearson Lemma - Proof

Neyman-Pearson Lemma - Proof

UMP Test - Normal Mean with Known Variance

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$. Find a UMP test of

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1 \;\; (\theta_1 > \theta_0).$$

UMP Test - Normal Mean with Known Variance

UMP Test - Normal Mean with Known Variance

UMP Test - Exponential

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim}$ Exponential(θ) and suppose we want to test

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1 \;\; (\theta_1 > \theta_0).$$

Find a UMP level α test.

UMP Test - Exponential

UMP Test - Exponential

UMP Test and Sufficient Statistics

Theorem

Consider testing

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta = \theta_1$$

and suppose $T(\boldsymbol{X})$ *is a sufficient statistic for* θ *and* $g(t|\theta)$ *is the pdf or pmf of* $T(\boldsymbol{X})$. *Then any test based on* T *satisfying*

$$\phi(t) = \begin{cases} 1, & g(t|\theta_1) > kg(t|\theta_0) \\ 0, & g(t|\theta_1) < kg(t|\theta_0) \end{cases} \quad (1)$$

for some $k > 0$ *such that* $E_{\theta_0}[\phi(T)] = \alpha$ *(2) is a UMP level* α *test. Moreover, if* ϕ *is a MP level* α *test, then* ϕ *satisfies (1) and (2).*

UMP Test - Binomial Proportion

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$, $0 < p < 1$, and suppose we want to test

$$H_0 : p = p_0 \quad \text{vs} \quad H_1 : p = p_1 \;\; (p_1 > p_0).$$

Find a UMP level α test.

UMP Test - Binomial Proportion

UMP Test - Binomial Proportion

UMP Test - Uniform

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$ and suppose we want to test

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1 \ \ (\theta_1 > \theta_0).$$

Find a UMP level α test.

UMP Test - Uniform

UMP Test - Uniform

UMP Test - Monotone Likelihood Ratio

Definition

A family of pdfs $\{g(t|\theta) : \theta \in \Theta\}$ for a univariate random variable T with real valued parameter θ has a monotone likelihood ratio (MLR) if for any $\theta_2 > \theta_1$ $(\theta_1, \theta_2 \in \Theta)$,

$$\frac{g(t|\theta_2)}{g(t|\theta_1)} \quad \text{is a non-decreasing function in } t$$

on the set $\{t : g(t|\theta_1) > 0 \text{ or } g(t|\theta_2) > 0\}$ where $c/0$ is defined to be ∞ if $c > 0$.

1. If the likelihood ratio is a non-increasing function in t then consider the family of densities for $-T$ to get a non-decreasing MLR.
2. Many common families of distributions have an MLR such as the normal distribution (with known variance) and unknown mean), Poisson, and binomial.
3. Any regular exponential family with $g(t|\theta) = h(t)c(\theta)e^{w(\theta)t}$ has an MLR if $w(\theta)$ is a non-decreasing function of θ.

UMP Test - Monotone Likelihood Ratio

Theorem

(Karlin-Rubin Theorem) Consider testing

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

Suppose that $T(\boldsymbol{X})$ is a sufficient statistic for θ and the family of pdfs (or pmfs) $\{g(t|\theta) : \theta \in \Theta\}$ of T has a non-decreasing MLR then

$$\phi(t) = \begin{cases} 1, & t > t_0 \\ \delta, & t = t_0 \\ 0, & t < t_0 \end{cases} \qquad (1)$$

is a UMP level α test where $E_{\theta_0}[\phi(T)] = \alpha$.

Proof - Karlin-Rubin Theorem

Proof - Karlin-Rubin Theorem

Proof - Karlin-Rubin Theorem

Karlin-Rubin Theorem - Other Cases

- If $g(t|\theta)$ is a non-**increasing** MLR, then a UMP level α test of
$$H_0 : \theta \leq \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0$$
is
$$\phi(t) = \begin{cases} 1, & t < t_0 \\ \delta, & t = t_0 \\ 0, & t \geq t_0 \end{cases}$$
with $E_{\theta_0}[\phi(T)] = \alpha$.

- If $g(t|\theta)$ is a non-decreasing MLR, then a UMP level α test of
$$H_0 : \theta \geq \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0$$
is
$$\phi(t) = \begin{cases} 1, & t < t_0 \\ \delta, & t = t_0 \\ 0, & t \geq t_0 \end{cases}$$
with $E_{\theta_0}[\phi(T)] = \alpha$.

Karlin-Rubin Theorem - Other Cases

- If $g(t|\theta)$ is a non-**increasing** MLR, then a UMP level α test of

$$H_0 : \theta \geq \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0$$

is

$$\phi(t) = \begin{cases} 1, & t > t_0 \\ \delta, & t = t_0 \\ 0, & t \leq t_0 \end{cases}$$

with $E_{\theta_0}[\phi(T)] = \alpha$.

Karlin-Rubin - Normal Mean with Known Variance

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$ and suppose we want to test

$$H_0 : \theta \geq \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0.$$

Find a UMP level α test.

Karlin-Rubin - Normal Mean with Known Variance

UMP Test Non-Existence

Let $X_1, X_2, \ldots, X_n \stackrel{iid}{\sim} N(\theta, 1)$ and consider testing

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

In this case, a UMP level α test does not exist. To see that, we will break the test into two one-sided tests.

UMP Test Non-Existence

UMP Test Non-Existence

UMPU Test

Definition

A test with power function $\beta(\theta)$ is **unbiased** if $\beta(\theta) > \beta(\theta')$ for all $\theta \in \Theta_1$ and $\theta' \in \Theta_0$.

Definition

A UMPU level α test is a level α unbiased test ϕ such that for any other level α unbiased test ϕ'

$$E_\theta[\phi(\boldsymbol{X})] \geq E_\theta[\phi'(\boldsymbol{X})], \text{ for all } \theta \in \Theta_1.$$

Note: A UMP level α test is always an unbiased test. (Compare the UMP test to the test function $\phi = \alpha$.)

UMPU Test

Theorem

Let $(X_1, X_2, \ldots, X_n) \sim f(\boldsymbol{x}|\theta)$ *(joint density where)*

$$f(\boldsymbol{x}|\theta) = a(\theta)h(\boldsymbol{x})e^{w(\theta)t(\boldsymbol{x})}, \theta \in \mathbb{R}$$

such that $w(\theta)$ *is a strictly increasing function of* θ*. Then a UMPU test of*

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0$$

is given by

$$\phi(t(\boldsymbol{x})) = \begin{cases} 1, & t(\boldsymbol{x}) < c_1 \text{ or } t(\boldsymbol{x}) > c_2 \\ \delta_i, & t(\boldsymbol{x}) = c_i, i = 1, 2 \\ 0, & c_1 < t(\boldsymbol{x}) < c_2 \end{cases}$$

where δ_i *and* c_i*,* $i = 1, 2$ *are determined by* $E_{\theta_0}[\phi(t(\boldsymbol{X}))] = \alpha$ *and* $E_{\theta_0}[t(\boldsymbol{X})\phi(t(\boldsymbol{X}))] = \alpha E_{\theta_0}[t(\boldsymbol{X})]$

UMPU Test - Two Sided Test of Normal Mean

Let $X_1, X_2, \ldots, X_n \stackrel{iid}{\sim} N(\theta, 1)$ and consider testing

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

Find a UMPU level α test.

Note that the joint density is

$$f(\boldsymbol{x}|\theta) = (2\pi)^{-n/2} e^{-\frac{n}{2}\theta^2} e^{-\frac{1}{2}\sum_{i=1}^n x_i^2} e^{\theta \sum_{i=1}^n x_i}$$

where $w(\theta) = n\theta$ is strictly increasing in θ with $T(\boldsymbol{X}) = \bar{X} \sim N(\theta, 1/n)$. A UMPU level α test is then given by

$$\phi(t) = \begin{cases} 1, & t < c_1 \text{ or } t > c_2 \\ 0, & c_1 < t < c_2 \end{cases}$$

satisfying $E_{\theta_0}[\phi(T)] = \alpha$ and $E_{\theta_0}[T\phi(T)] = \alpha E_{\theta_0}[T] = \alpha\theta_0$.

UMPU Test - Two Sided Test of Normal Mean

To find the UMPU test, we would need to find a c_1 and a c_2 satisfying the following two equations:

1. $\alpha = P_{\theta_0}(T < c_1) + P_{\theta_0}(T > c_2)$
2. $\alpha\theta_0 = \int_{-\infty}^{c_1} t\varphi_{\theta_0,1/n}(t)\, dt + \int_{c_2}^{\infty} t\varphi_{\theta_0,1/n}(t)\, dt$

where $\varphi_{\theta_0,1/n}$ is the density of a normal distribution with mean θ_0 and variance $1/n$. Solving these two equations yields the UMPU test

$$\phi(t) = \begin{cases} 1, & t < \theta_0 - z_{1-\alpha/2}/\sqrt{n} \text{ or } t > \theta_0 + z_{1-\alpha/2}/\sqrt{n} \\ 0, & \theta_0 - z_{1-\alpha/2}/\sqrt{n} < t < \theta_0 + z_{1-\alpha/2}/\sqrt{n} \end{cases}$$

Interval Estimation

Given data $\boldsymbol{X} = (X_1, X_2, \dots, X_n) \sim f(\boldsymbol{x}|\theta)$, out goal is to find a region of plausible values for the parameter θ.

Definition

An interval estimate of a real valued parameter θ is any pair of function $L(\boldsymbol{X}) = L(X_1, X_2, \dots, X_n)$ and $U(\boldsymbol{X}) = U(X_1, X_2, \dots, X_n)$ of a sample $\boldsymbol{X}$ that satisfies $L(\boldsymbol{X}) \leq U(\boldsymbol{X})$ for every sample $\boldsymbol{X} \in \mathcal{X}$. If $\boldsymbol{x}$ is observed the inference $L(\boldsymbol{x}) \leq \theta \leq U(\boldsymbol{x})$ is made and we write the interval estimator of θ as $[L(\boldsymbol{x}), U(\boldsymbol{x})]$.

Note: If $L(\boldsymbol{x}) = -\infty$, then $\theta \leq U(\boldsymbol{x})$ is an upper bound. If $U(\boldsymbol{x}) = \infty$, then $L(\boldsymbol{x}) \leq \theta$. Both are called one-sided interval estimators.

Interval Estimation

Let $X_1, X_2, X_3, X_4 \overset{iid}{\sim} N(\mu, 1)$. Assume an interval estimator for μ is given by $[\bar{X} - 1, \bar{X} + 1]$. Then

$$\begin{aligned}
P_\mu(\theta \in [\bar{X} - 1, \bar{X} + 1]) &= P_\mu(\bar{X} - 1 \leq \mu \leq \bar{X} + 1) \\
&= P_\mu(-\mu - 1 \leq -\bar{X} \leq -\mu + 1) \\
&= P_\mu(\mu + 1 > \bar{X} > \mu - 1) \\
&= P_\mu\left(\frac{\mu - 1 - \mu}{1/2} \leq Z \leq \frac{\mu + 1 - \mu}{1/2}\right) \\
&= P_\mu(-2 \leq Z \leq 2) \approx 0.95
\end{aligned}$$

Interval Estimation - Coverage Probability

Definition

For an interval estimator $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ of a parameter θ, the **coverage probability** of $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ is the probability that the random interval $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ covers the true parameter θ.

Coverage Probability of $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ at $\theta = P_\theta(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})])$.

Definition

For an interval estimator $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ of a parameter θ, the **confidence coefficient** of $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ is

$$\inf_{\theta \in \Theta} P_\theta(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})]).$$

Interval Estimation - Coverage Probability

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$. Let's consider two interval estimators of $\theta \in \Theta = (0, \infty)$

(a) $[aX_{(n)}, bX_{(n)}]$ for some $1 \leq a < b$

(b) $[X_{(n)} + c, X_{(n)} + d]$ for some $0 \leq c < d$

Methods of Finding Interval Estimators - Inverting a Test

Consider size α test of

$$H_0 : \theta = \theta_0$$

with a rejection region $\mathcal{R}$. Consider the acceptance region $A(\theta_0) = \mathcal{R}^c$. Then

$$1 - \alpha = P_{\theta_0}(\boldsymbol{X} \in A(\theta_0)).$$

Thus if we can invert this region, then we have an interval estimate with coverage probability $1 - \alpha$.

Methods of Finding Interval Estimators - Inverting a Test

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$. We want a confidence interval for θ. Consider the test of

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

The UMPU test based on $T(\boldsymbol{X}) = \bar{X}$ is given by

$$\phi(t) = \begin{cases} 1, & t < \theta_0 - z_{1-\alpha/2}/\sqrt{n} \text{ or } t > \theta_0 + z_{1-\alpha/2}/\sqrt{n} \\ 0, & \theta_0 - z_{1-\alpha/2}/\sqrt{n} < t < \theta_0 + z_{1-\alpha/2}/\sqrt{n} \end{cases}$$

Methods of Finding Interval Estimators - Inverting a Test

Methods of Finding Interval Estimators - Inverting a Test

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$. Suppose we want a one sided confidence interval. We can construct such an interval by inverting the acceptance region of

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0$$

We found the following UMP level α test based in $T(\boldsymbol{X}) = \bar{X}$

$$\phi(t) = \begin{cases} 1, & t < \theta_0 - z_{1-\alpha/2}/\sqrt{n} \\ 0, & \text{otherwise} \end{cases}.$$

Methods of Finding Interval Estimators - Inverting a Test

Methods of Finding Interval Estimators - Inverting a LRT

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim}$ Exponential(θ) and suppose we want a two sided confidence interval for θ. Consider the two sided LRT of

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

Methods of Finding Interval Estimators - Inverting a LRT

Methods of Finding Interval Estimators - Inverting a LRT

Methods of Finding Interval Estimators - Pivotal Quantities

In the previous example, we based our confidence on $T(\boldsymbol{X}, \theta) = \frac{2\sum_{i=1}^{n} X_i}{\theta}$ and the distribution of $T(\boldsymbol{X}, \theta)$ does not depend on θ. $T(\boldsymbol{X}, \theta)$ is called a pivot.

Definition

A random variable $T(\boldsymbol{X}, \theta) = T(X_1, X_2, \ldots, X_n, \theta)$ is a **pivotal quantity** or pivot if the distribution of $T(\boldsymbol{X}, \theta)$ is independent of all parameters. That is, if $\boldsymbol{X} \sim f(\boldsymbol{x}|\theta)$, then the distribution of $T(\boldsymbol{X}, \theta)$ is independent of θ for all θ.

Note: Once we have a pivot $T(\boldsymbol{X}, \theta)$ for some given α, we let

$$P_\theta(a \leq T(\boldsymbol{X}, \theta) \leq b) = 1 - \alpha$$

define a $1 - \alpha$ confidence interval for θ by solving for θ in terms of $\boldsymbol{X}$.

Methods of Finding Interval Estimators - Pivotal Quantities

1. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$. Recall that $T(\boldsymbol{X}) = X_{(n)}$ is sufficient with pdf
$$g(t|\theta) = \frac{nt^{n-1}}{\theta^n}, 0 < t < \theta.$$
Let $U(\boldsymbol{X}, \theta) = X_{(n)}/\theta$, then U has pdf
$$g(u) = nu^{n-1}, 0 < u < 1.$$
The pdf of U does not depend on θ, so $X_{(n)}/\theta$ is a pivot.
2. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\theta)$. Then $\sum_{i=1}^n X_i \sim \text{Gamma}(n, \theta)$ is sufficient for θ. and $\sum_{i=1}^n X_i/\theta \sim \text{Gamma}(n, 1)$ does not depend on θ.
3. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$. Recall that $\bar{X}$ is sufficient for θ and note that
$$\sqrt{n}(\bar{X} - \theta) = \frac{\bar{X} - \theta}{1/\sqrt{n}} \sim N(0, 1).$$

Methods of Finding Interval Estimators - Pivotal Quantities

(Location-Scale Pivots) Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f$ given by

Form of pdf	Type of pdf	Pivot
$f(x-\mu)$	Location	$\bar{X}-\mu$
$\frac{1}{\sigma}f\left(\frac{x}{\sigma}\right)$	Scale	$\frac{\bar{X}}{\sigma}$
$\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$	Location-Scale	$\frac{\bar{X}-\mu}{S}$

Homework: Show that each of these are pivotal quantities

Methods of Finding Interval Estimators - Pivotal Quantities

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$. Find a pivotal quantity.

Methods of Finding Interval Estimators - Pivotal Quantities

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$ with pdf

$$f(x) = (2\pi)^{-1/2} e^{-(x-\theta)^2/2}.$$

Find a pivotal quantity.

Methods of Finding Interval Estimators - Pivotal Quantities

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$ with pdf

$$f(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\theta)^2/2\sigma^2}.$$

Find a confidence interval for θ and for σ^2.

Methods of Finding Interval Estimators - Pivotal Quantities

Methods of Finding Interval Estimators - Pivotal Quantities

Methods of Finding Interval Estimators - Pivotal Quantities

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f(x|\mu) = e^{-(x-\mu)}, x > \mu$. Then $X_i - \mu \overset{iid}{\sim}$ Exp(1), so we have the following pivots

1. $\bar{X} - \mu = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \sim \text{Gamma}(n, 1/n)$
2. $\sum_{i=1}^{n}(X_i - \mu) \sim \text{Gamma}(n, 1)$
3. $X_{(1)} - \mu \sim \text{Exp}(1/n)$

Evaluating Interval Estimates - Shortest Interval Length

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$. We have already seen that a pivot is $\sqrt{n}(\bar{X} - \mu) \sim N(0, 1)$. We want to choose an $a < b$ such that

$$1 - \alpha = P(a \leq \sqrt{n}(\bar{X} - \mu) \leq b)$$

which results in a $(1 - \alpha) \times 100\%$ confidence interval

$$\left[\bar{X} - b\frac{1}{\sqrt{n}}, \bar{X} - a\frac{1}{\sqrt{n}}\right].$$

Consider $\alpha = 0.1$ and the following choices of a and b

a	b	$P(Z < a)$	$P(Z > b)$	$b - a$
-1.34	2.33	0.09	0.01	3.67
-1.44	1.96	0.075	0.025	3.40
-1.65	1.65	0.05	0.05	3.30

The length of the confidence interval in each case is $\frac{b - a}{\sqrt{n}}$. The shortest interval corresponds to $a = z_{\alpha/2} = -1.65$ and $b = z_{1-\alpha/2} = 1.65$

Evaluating Interval Estimates - Shortest Interval Length

Theorem

Let $f(x)$ be a unimodal pdf. If the interval $[a, b]$ satisfies

1. $\int_a^b f(x)dx = 1 - \alpha$
2. $f(a) = f(b) > 0$
3. $a \leq x^\star \leq b$, *where* $x^\star$ *is a mode of* $f(x)$

then $[a, b]$ is the shortest interval among all intervals that satisfy (1).

1. If $f(x)$ is also symmetric about the y-axis (such as the standard normal and Student t's distributions), then a, b will be such that $P(X < a) = P(X > b) = \alpha/2$
2. If $f(x)$ is strictly increasing over a finite interval $[\gamma, \beta]$, then the shortest interval is $[a, b]$ where $b = \beta$ and $\int_\gamma^a f(x)dx = \alpha$.
3. Similarly, if $f(x)$ is strictly decreasing over a finite interval $[\gamma, \beta]$, then the shortest interval is $[a, b]$ where $a = \gamma$ and $\int_b^\beta f(x)dx = \alpha$.

Evaluating Interval Estimates - Shortest Interval Length

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$. Consider the pivot $X_{(n)}/\theta$ with pdf

$$g(u) = nu^{n-1}, 0 < u < 1.$$

Find the shortest length $100(1-\alpha)\%$ confidence interval for θ.

Evaluating Interval Estimates - Shortest Interval Length

Bayesian Inference

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ where $\boldsymbol{X}|\theta \sim f(\boldsymbol{x}|\theta)$, $\theta \in \Theta$. In the Bayesian paradigm, θ is treated as a random variable and assigned a prior distribution $\theta \sim \pi(\theta)$. Inference is then based on the posterior distribution of $\theta|\boldsymbol{X}$,

$$\pi(\theta|\boldsymbol{X}) = \frac{f(x,\theta)}{\int_{-\infty}^{\infty} f(x,\theta)\, d\theta} = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{\int_{-\infty}^{\infty} f(x,\theta)\, d\theta}.$$

If we observe $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ given θ, we update the distribution of θ base on the observed data $\boldsymbol{X}$.

In this section, we will introduce Bayesian interval estimation and hypothesis testing.

Bayesian Interval Estimation

Definition

A $100(1-\alpha)\%$ credible set for θ is a subset $S_{\boldsymbol{x}} \subseteq \Theta$ such that $P(\theta \in S_{bx}|\boldsymbol{x}) = 1-\alpha$.

Like in frequentist CI, a common way to build a a credible interval is to find the $\alpha/2\%$ and $(1-\alpha/2)\%$ quantiles, say ξ_1 and ξ_2 respectively, of the posterior distribution to form the $100(1-\alpha)\%$ credible interval (ξ_1, ξ_2).

Definition

A $100(1-\alpha)\%$ highest posterior density (HPD) set for θ is a subset $S_{\boldsymbol{x}} = \{\theta \in \Theta : \pi(\theta|\boldsymbol{x}) \geq c_\alpha\}$ such that $P(\theta \in S_{bx}|\boldsymbol{x}) = 1-\alpha$.

The HPD set has the smallest size (or length) among all $100(1-\alpha)\%$ credible sets. If $\pi(\theta|\boldsymbol{x})$ is unimodal and continuous, it is the interval satisfying $\pi(a|\boldsymbol{x}) = \pi(b|\boldsymbol{x}) = c_\alpha$.

Bayesian Interval Estimation

Let $X_1, X_2, \ldots, X_n|\theta \overset{iid}{\sim} N(\theta, \sigma^2)$, σ^2 known. Using the prior $\theta \sim N(\mu, \tau^2)$, we showed earlier that posterior distribution of $\theta|\boldsymbol{X}$ is

$$\theta|\boldsymbol{X} \sim N\left(\frac{n\tau^2\bar{x} + \mu\sigma^2}{\sigma^2 + n\tau^2}, \frac{\sigma^2\tau^2}{n\tau^2 + \sigma^2}\right).$$

Find a $100(1-\alpha)\%$ credible interval for θ.

Bayesian Hypothesis Testing

Suppose we'd like to test $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$, where $\Theta_0 \cap \Theta_1 = \emptyset$ and $\Theta_0 \cup \Theta_1 = \Theta$. For hypothesis testing, it's often more convenient to define the prior as a mixture distribution. Let $\pi_0 = \mathbb{P}(\theta \in \Theta_0)$ and $\pi_1 = \mathbb{P}(\theta \in \Theta_1) = 1 - \pi_0$. Then given $\theta \in \Theta_0$, define a distribution $p_0(\theta)$ on Θ_0. Similarly, define $p_1(\theta)$ on Θ_1. The overall prior is then

$$\pi(\theta) = \pi_0 p_0(\theta)\mathbf{1}(\theta \in \Theta_0) + \pi_1 p_1(\theta)\mathbf{1}(\theta \in \Theta_1).$$

The evidence for H_1 over H_0 is measured by the posterior odds

$$\begin{aligned} \frac{\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x})}{\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x})} &= \frac{\int_{\Theta_1} \pi(\theta|\boldsymbol{x})d\theta}{\int_{\Theta_0} \pi(\theta|\boldsymbol{x})d\theta} \\ &= \frac{\int_{\Theta_1} \pi_1 p_1(\theta)f(\boldsymbol{x}|\theta)d\theta}{\int_{\Theta_0} \pi_0 p_0(\theta)f(\boldsymbol{x}|\theta)d\theta} \\ &= \frac{1-\pi_0}{\pi_0}\frac{\int_{\Theta_1} p_1(\theta)f(\boldsymbol{x}|\theta)d\theta}{\int_{\Theta_0} p_0(\theta)f(\boldsymbol{x}|\theta)d\theta} \end{aligned}$$

Bayesian Hypothesis Testing

For simple hypothesis $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$, as $\mathbb{P}(\theta = \theta_k | \boldsymbol{x}) \propto \pi_k f(\boldsymbol{x}|\theta_k)$, $k = 0, 1$, we simply have

$$\frac{\mathbb{P}(\theta = \theta_1|\boldsymbol{x})}{\mathbb{P}(\theta = \theta_0|\boldsymbol{x})} = \frac{1 - \pi_0}{\pi_0} \frac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)}.$$

This equation can also be derived from the general formula. Noting that $\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \{\theta_1\}$, and $p_0(\theta_0) = p_1(\theta_1) = 1$ (degenerate with probability one at a single point), we have $\int_{\Theta_k} p_k(\theta) f(\boldsymbol{x}|\theta) d\theta = \int_{\Theta_k} f(\boldsymbol{x}|\theta) dP_k(\theta) = f(\boldsymbol{x}|\theta_k)$, $k = 0, 1$, where $P_k(\theta)$ is a step function with a jump from 0 to 1 at θ_k (counting measure).\

A reasonable test is to rejected H_0 if $\frac{\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x})}{\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x})} > C$. For simple hypothesis, if $\pi_0 = \pi_1 = 0.5$, then the rejection region simplifies to $\frac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|\theta_0)} > C$, the same as the likelihood ratio test, which is the MP test.

Bayesian Hypothesis Testing

How to choose the threshold C? We have two decisions (actions):

- Reject H_0. If H_1 is true (which happens with probability $\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x})$), there is no penalty. If H_0 is true (which happens with probability $\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x})$), the penalty is K_0. Then the expected posterior loss is

$$K_0\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x}) + 0\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x}) = K_0\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x}).$$

- Accept H_0. If H_1 is true, there is a penalty of K_1. If H_0 is true, there is no penalty. Then the expected posterior loss is $K_1\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x})$

To minimize the expected posterior loss, we choose "reject H_0" when $K_0\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x}) < K_1\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x})$, or equivalently, $\frac{K_0}{K_1} < \frac{\mathbb{P}(\theta \in \Theta_1|\boldsymbol{x})}{\mathbb{P}(\theta \in \Theta_0|\boldsymbol{x})}$. Consequently, $C = K_0/K_1$. If $K_0 = K_1$, $C = 1$.

Bayesian Hypothesis Testing

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, where σ is known. Consider testing $H_0 : \mu = \mu_0$ vs. $H_1 : \mu = \mu_1$ $(\mu_1 > \mu_0)$. Let $\pi_0 = \mathbb{P}(\mu = \mu_0)$ and $\pi_1 = \mathbb{P}(\mu = \mu_1)$ be the prior distribution.

Asymptotic Inference

Suppose that $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$, then $\hat{\theta} \pm z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ is a $(1 - \alpha) \times 100\%$ asymptotic CI for θ. If in addition $\hat{\sigma}_n \xrightarrow{P} \sigma$, then

$$\sqrt{n}(\hat{\theta} - \theta)/\hat{\sigma}_n = \sqrt{n}(\hat{\theta} - \theta)/\sigma \times (\sigma/\hat{\sigma}_n) \xrightarrow{d} \mathcal{N}(0, 1).$$

Thus for large n

$$\hat{\theta} \sim \mathcal{N}(\theta, \hat{\sigma}_n^2/n),$$

and we can construct approximate (asymptotically valid) CI and HT as we did in the normal case, i.e $\hat{\theta} \pm z_{1-\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}}$ is also a $(1 - \alpha) \times 100\%$ asymptotic CI for θ.

Asymptotic Inference

Suppose that $X_1, \ldots, X_n$ are i.i.d. with unknown mean μ and unknown variance $\sigma^2 < \infty$. Find a CI for μ.

Asymptotic Inference

Suppose $X_1, \ldots, X_n \overset{i.i.d.}{\sim}$ Bernoulli(p). Find an asymptotic CI for p.