# 3. TESTING HYPOTHESES

**3**.1 Hypothesis testing: basic setting

Statistical estimation and tests address *different kind* of practical problems.

We work with a family of distributions $f(\mathbf{x}; \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \Theta$, and $f(\mathbf{x}; \boldsymbol{\theta})$ is completely known apart from $\boldsymbol{\theta}$.

**Basic setting**:
A statistical test make a decision between two sets of hypotheses on unknown parameter $\boldsymbol{\theta}$, namely

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

where $\Theta_0, \Theta_1$ are subsets of $\Theta$. We always assume

$$\Theta_0 \cap \Theta_1 = \emptyset \text{ (empty set)},$$

i.e. $H_0$ and $H_1$ are distinct, mutually exclusive, although we do not always require $\Theta_0 \cup \Theta_1 = \Theta$.

Based on a sample $\mathbf{X} = (X_1, \cdots, X_n)^\tau$, a statistical test is to make a decision either

**t**o reject $H_0$, or
**n**ot to reject $H_0$.

In latter case, we claim that there is **n**o significant difference between the null hypothesis and the underlying distribution according to the observed data.

**R**emark. The setting is **n**ot symmetric in the sense that the two hypotheses $H_0$ and $H_1$ will be treated differently.

Some specific terms used statistical tests are now in order:

- A *Hypothesis* is an assumption made about the value of $\boldsymbol{\theta}$.
- A *Simple Hypothesis* completely specifies $\boldsymbol{\theta}$ to a known constant such as H: $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.
- A *Composite Hypothesis* is any hypothesis that is not simple. For instance the hypothesis H: $\boldsymbol{\theta} \in [1, \infty)$.
- A *Test Statistic* is a statistic based on which the decision is made.
- A *Null Hypothesis*, usually denoted as $H_0$, is the basic or default hypothesis.
- The *Alternative Hypothesis*, usually denoted as $H_1$, is a hypothesis that is to be compared with the Null Hypothesis.

Mathematically speaking, a statistical test is equivalent to define a binary function

$$\phi(\mathbf{X}) = \begin{cases} 1 & \text{when reject } H_0, \\ 0 & \text{when not reject } H_0. \end{cases}$$

$\phi(\cdot)$ is called a **d**ecision rule, which practically divides the sample space of $\mathbf{X}$ into two subspaces: *rejection region* (which is also called '**c**ritical region') and its compliment which is often **imprecisely** labelled as *'acceptance' region.*

There are two types of errors that can be made in a statistical test, which are displayed in the table below.

| | | Decision Made | |
|---|---|---|---|
| | | $H_0$ not rejected $\Theta_0$ | $H_0$ rejected $\Theta_1$ |
| True State of Nature | $H_0$ $\Theta_0$ | Correct Decision | Type I Error |
| | $H_1$ $\Theta_1$ | Type II Error | Correct Decision |

Ideally we would like to have a test that minimises the probabilities of making both types of errors, which unfortunately is not feasible.

**Construction of a statistical test**:
we control the probability of Type I error under a given level, say, $\alpha$, and then we minimize the probability of Type II error.

**D**efinition. A test is said to have *significance level* $\alpha$ if

$$\sup_{\boldsymbol{\theta} \in \Theta_0} P_{\boldsymbol{\theta}}(H_0 \text{ is rejected}) \leq \alpha.$$

**D**efinition. A test $T$ is said to be of *size* $\alpha$ if

$$\sup_{\boldsymbol{\theta} \in \Theta_0} P_{\boldsymbol{\theta}}(H_0 \text{ is rejected}) = \alpha.$$

**R**emark. (i) In practice, we usually take the significance level $\alpha = 0.1, 0.05$ or $0.01$.
(ii) The size of the test is no greater than the significance level.

**D**efinition. The *power function* of a test is defined as

$$\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(H_0 \text{ is rejected}), \quad \boldsymbol{\theta} \in \Theta_1.$$

In practice, it is useful to extend $\beta(\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \Theta$. For $\boldsymbol{\theta} \in \Theta_0$, $\beta(\boldsymbol{\theta})$ must be smaller than the prescribed significance level, and is the probability of making a Type I error at the parameter value $\boldsymbol{\theta}$. For $\boldsymbol{\theta} \in \Theta_1$, $1 - \beta(\boldsymbol{\theta})$ is the probability of making a Type II error.

If a test is defined in terms of decision rule $\phi$, then

$$\beta(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}[\phi(\mathbf{X})],$$

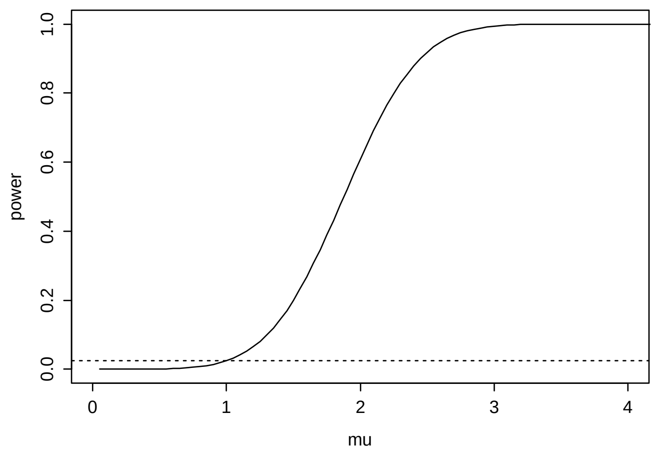
because

$$\begin{aligned} E[\phi(\mathbf{X})] &= \int \phi(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}) d\mathbf{x} \\ &= P_{\boldsymbol{\theta}}\{\phi(\mathbf{X}) = 1\} = P_{\boldsymbol{\theta}}\{H_0 \text{ is rejected}). \end{aligned}$$

**E**xample 1. Suppose that we have a random sample of size 5 from a normal distribution with mean $\mu$ and variance 1. Let us consider:

$$\begin{aligned} H_0&: \quad \mu = 1 \\ H_1&: \quad \mu > 1 \end{aligned}$$

One possible test statistic is $\bar{X}$. Intuitively we would reject $H_0$ for large values of $\bar{X}$, say, $\bar{X} > K$. The constant $K$ is determined by the prescribed significance level $\alpha$. Let $\alpha = 0.025$. Since $\sqrt{5}(\bar{X} - 1) \sim N(0, 1)$, $K = 1 + 1.96/\sqrt{5}$. We reject $H_0$ if $\bar{X} > 1 + 1.96/\sqrt{5}$. The power function is shown below.




The power is small for $\mu = 1$, and rises quite fast as $\mu$ increases past 1.

**Question**: What should we do if we change the alternative hypothesis to $H_1 : \mu < 1$ or $H_1 : \mu \neq 1$.

The above null hypothesis is simple, because $\Theta_0 = \{1\}$, but the alternative hypothesis is composite because $\Theta_1 = (1, \infty)$. If instead we used the Null Hypothesis $H_0$: $\mu \leq 1$, that would be a composite null hypothesis. It is convenient to have one value to summarise the probability of Type I error, even when the hull hypothesis is composite.

**Question**: Is the test constructed above still reasonable for testing

$$H_0 : \ \mu \leq 1 \quad \text{vs} \quad H_1 : \mu > 1\ ?$$

If so, at which value of $\mu$ under $H_0$ does the probability of type I error obtain the maximum?

**3**.2 Neyman-Pearson Lemma for testing simple hypotheses

**Most powerful test (MPT)**: among all the tests of size $\alpha$, the one with maximum power $\beta(\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \Theta_1$ is called the MPT. When $\Theta_1$ consists of more than one point, it is often called the uniformly MPT (UMPT).

**N**ote. UMPT may not exist.

Let $L(\boldsymbol{\theta}; \mathbf{X})$ be the likelihood function with observation $\mathbf{X}$.

**T**heorem 1. (*Neyman-Pearson Lemma*)
The MPT for hypotheses

$$H_0 : \ \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{against} \quad H_1 : \ \boldsymbol{\theta} = \boldsymbol{\theta}_1$$

rejects $H_0$ when

$$L(\boldsymbol{\theta}_1; \mathbf{X}) > K L(\boldsymbol{\theta}_0; \mathbf{X}),$$

and does not reject $H_0$ when

$$L(\boldsymbol{\theta}_1; \mathbf{X}) < K L(\boldsymbol{\theta}_0; \mathbf{X}),$$

where constant $K$ is determined by the size of the test.

**R**emark. In general the MPT rejects $H_0$ if the **l**ikelihood ratio

$$LR(\mathbf{X}) = \frac{L(\boldsymbol{\theta}_1; \mathbf{X})}{L(\boldsymbol{\theta}_0; \mathbf{X})} > K,$$

which indicates that $\boldsymbol{\theta}_1$ is relatively favoured over $\boldsymbol{\theta}_0$ according to the likelihood function.

**P**roof. For each test $T$, we define the decision rule $\phi_T(\mathbf{x})$ as follows

$$\phi_T(\mathbf{X}) = \begin{cases} 1 & \text{Reject } H_0 \\ 0 & \text{Do not Reject } H_0. \end{cases}$$

The power $\beta(\boldsymbol{\theta})$ is equal to $E\phi_T(\mathbf{X})$.

Suppose that $T$ is a test of size $\alpha$ that satisfies the conditions of the Neyman-Pearson Lemma, and $S$ is some other test of size $\alpha$. Then

$$\begin{aligned}
\beta_S(\boldsymbol{\theta}_1) = E\phi_S(\mathbf{x}) &= \int \phi_S(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}_1) d\mathbf{x} \\
&= \int \phi_S(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}_1) d\mathbf{x} - K \int \phi_S(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}_0) d\mathbf{x} + K\alpha \\
&= \int \phi_S(\mathbf{x}) [f(\mathbf{x}; \boldsymbol{\theta}_1) - K f(\mathbf{x}; \boldsymbol{\theta}_0)] d\mathbf{x} + K\alpha \\
&= \int \phi_S(\mathbf{x}) [L(\boldsymbol{\theta}_1; \mathbf{x}) - K L(\boldsymbol{\theta}_0; \mathbf{x})] d\mathbf{x} + K\alpha \\
&\leq \int_{\{L(\boldsymbol{\theta}_1; \mathbf{x}) > K L(\boldsymbol{\theta}_0; \mathbf{x})\}} \phi_S(\mathbf{x}) [L(\boldsymbol{\theta}_1; \mathbf{x}) - K L(\boldsymbol{\theta}_0; \mathbf{x})] d\mathbf{x} + K\alpha \\
&\leq \int \phi_T(\mathbf{x}) [L(\boldsymbol{\theta}_1; \mathbf{x}) - K L(\boldsymbol{\theta}_0; \mathbf{x})] d\mathbf{x} + K\alpha \\
&= \int \phi_T(\mathbf{x}) [f(\mathbf{x}; \boldsymbol{\theta}_1) - K f(\mathbf{x}; \boldsymbol{\theta}_0)] d\mathbf{x} + K\alpha \\
&= \int \phi_T(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}_1) d\mathbf{x} - K \int \phi_T f(\mathbf{x}; \boldsymbol{\theta}_0) d\mathbf{x} + K\alpha \\
&= \beta_T(\boldsymbol{\theta}_1) - K\alpha + K\alpha \\
&= \beta_T(\boldsymbol{\theta}_1).
\end{aligned}$$

**E**xample 2. Suppose that $X_1, X_2, \ldots, X_n$ are a random sample from $N(\mu, \sigma^2)$, where we treat $\sigma^2$ as a known quantity. Suppose we want to test:

$$H_0 : \mu = 0 \quad \text{against} \quad H_1 : \mu = 5.$$

The likelihood ratio is

$$\begin{aligned}
LR &= \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\sum_{i=1}^n (X_i - 5)^2/(2\sigma^2)\right)}{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\sum_{i=1}^n X_i^2/(2\sigma^2)\right)} \\
&= \exp\left(5n(2\bar{X} - 5)/(2\sigma^2)\right).
\end{aligned}$$

We can look for a most powerful test by noting that $LR > K$ is equivalent to $\bar{X} > K_1$, and *the critical value* $K_1$ is determined by the size of the test. So, if we find a test of this form which has size $\alpha$, it will be a most powerful test of size $\alpha$. It is easy to see that the test that rejects $H_0$ iff $\bar{X} > z_\alpha \sigma/\sqrt{n}$ is of this form and is therefore a most powerful test of size $\alpha$.

**Question**: If we change the alternative hypothesis to $H_1 : \mu = 50$, what is the MPT then?

**E**xample 3. Let $(X_1, \cdots, X_n)$ be a sample from Poisson distribution with mean $\mu$. To test

$$H_0: \ \mu = 2 \quad \text{against} \quad H_1: \ \mu = 6,$$

the likelihood ratio is

$$LR = \frac{6^{\sum X_i} e^{-6n}/(X_1! \cdots X_n!)}{2^{\sum X_i} e^{-2n}/(X_1! \cdots X_n!)} = 3^{\sum X_i} e^{-4n}.$$

According to the N-P lemma, the MPT will reject $H_0$ if $LR > K$, or equivalently, $\sum_i X_i > K_1$, where the critical value $K_1$ is determined by the significance level of the test.

Suppose $n = 4$, then $\sum_{i=1}^4 X_i \sim Poisson(8)$ under $H_0$. From the table for Poisson CDFs, the size of test $\alpha = 0.064$ with $K_1 = 12$, and $\alpha = 0.034$ with $K_1 = 13$.

When $n = 8$, $\sum_{i=1}^8 X_i \sim Poisson(16)$ under $H_0$. The size of test $\alpha = 0.058$ with $K_1 = 22$, and $\alpha = 0.037$ with $K_1 = 23$.

**3**.3 Uniformly Most Powerful Tests

Although the Neyman-Pearson Lemma was designed for testing simple hypotheses, it is possible to use it to construct the UMPT for, for example, *one-sided null hypothesis* against *one-sided alternative*.

**A** general setting: For hypotheses

$$H_0: \ \theta \in \Theta_0 \quad \text{vs} \quad H_1: \ \theta \in \Theta_1, \tag{4}$$

$\phi(\mathbf{X})$ is the decision rule of a test at significant level $\alpha$, i.e.

$$E_\theta\{\phi(\mathbf{X})\} \le \alpha \qquad \text{for all } \theta \in \Theta_0.$$

If $\phi(\mathbf{X})$ is also the MPT of size $\alpha$ for *simple hypotheses*

$$H_0: \ \theta = \theta_0 \quad \text{vs} \quad H_1: \ \theta = \theta_1$$

for **some** $\theta_0 \in \Theta_0$ and **all** $\theta_1 \in \Theta_1$. Then $\phi(\mathbf{X})$ is the UMPT for hypotheses (4).

We look at a simple scenario first — *UMPTs for simple* $H_0$ *and one-sided* $H_1$.

Suppose the MPT for simple hypotheses

$$H_0: \ \theta = \theta_0 \quad \text{against} \quad H_1: \ \theta = \theta_1$$

*does not change its form for all* $\theta_1 \in \Theta_1$. Then it is also the UMPT for

$$H_0: \ \theta = \theta_0 \quad \text{against} H_1: \ \theta \in \Theta_1.$$

Such an UMPT often exists for $\Theta_1 = (\theta_0, \infty)$ or $\Theta_1 = (-\infty, \theta_0)$.

**E**xample 4. If $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, 1)$, then the most powerful test for testing

$$H_0 : \mu = 1 \quad \text{vs} \quad H_1 : \mu = 2$$

is obtained by rejecting $H_0$ if and only if $\bar{X} > 1 + z_\alpha/\sqrt{n}$. The same test is obtained for

$$H_0 : \mu = 1 \quad \text{vs} \quad H_1 : \mu = \mu_1$$

for every $\mu_1 > 1$. So this test is uniformly most powerful for

$$H_0 : \mu = 1 \quad \text{vs} \quad H_1 : \mu > 1.$$

Although the UMPT does not depend on the value $\mu_1$ specified under $H_1$, its power varies over $\{\mu|\mu > 1\}$. In fact

$$\begin{aligned} \beta(\mu) &= P_\mu\{\bar{X} > 1 + z_\alpha/\sqrt{n}\} \\ &= P_\mu\{\sqrt{n}(\bar{X} - \mu) > \sqrt{n}(1 - \mu) + z_\alpha\} \\ &= 1 - \Phi\{z_\alpha - \sqrt{n}(\mu - 1)\} = \Phi\{\sqrt{n}(\mu - 1) - z_\alpha\}, \end{aligned}$$

where $\Phi(\cdot)$ is the CDF of N(0,1). Thus the power increases as $\mu$ increases.

Note that the test is not the UMPT for $H_1 : \mu \neq 1$.

**N**ote. The UMPT usually does not exist for two-sided (composite) alternative hypothesis.

**E**xample 5. Let $Y$ have the binomial distribution $Bin(n, p)$. Find the UMPT for testing

$$H_0 : p = p_0 \quad \text{vs} \quad H_1 : p > p_0.$$

Let $p_1 > p_0$. The LR for testing the $H_0$ above against $H_1 : p = p_1$ is

$$\begin{aligned} LR &= \frac{\binom{n}{Y} p_1^Y (1 - p_1)^{n-Y}}{\binom{n}{Y} p_0^Y (1 - p_0)^{n-Y}} \\ &= \left(\frac{p_1(1 - p_0)}{p_0(1 - p_1)}\right)^Y \left(\frac{1 - p_1}{1 - p_0}\right)^n \end{aligned}$$

Note that $p_1(1-p_0) > p_0(1-p_1)$ since $p_1 > p_0$. Thus $LR > K$ is equivalent to $Y > K_1$. Thus the UMPT rejects $H_0$ iff $Y > K_1$ with the size

$$P(Y > K_1|p = p_0).$$

**E**xample 6. Let $(X_1, \cdots, X_n)$ be a random sample from an exponential distribution with mean $1/\lambda$. We are interested in testing

$$H_0 : \ \lambda \leq \lambda_0 \quad \text{vs} \quad H_1 : \ \lambda > \lambda_0.$$

For

$$H_0 : \ \lambda = \lambda_0 \quad \text{vs} \quad H_1 : \ \lambda = \lambda_1,$$

the MPT rejects $H_0$ iff $\sum_{i=1}^n X_i \leq K$ for any $\lambda_1 > \lambda_0$, where $K$ is determined by

$$P_{\lambda_0}\{\sum_{i=1}^n X_i < K\} = \alpha.$$

It is easy to verify that for $\lambda < \lambda_0$,

$$P_\lambda\{\sum_{i=1}^n X_i < K\} < \alpha.$$

Hence the MPT for the simple null hypothesis against simple alternative is also the UMPT for the composite hypotheses.

**3**.4 Likelihood Ratio Tests

We now deal with the most popular ways of constructing tests when both null and alternative hypotheses are composite. There are no guaranteed optimality properties in small samples for these tests, but in regular cases they usually have good power for large sample sizes.

Let $\mathbf{X} \sim f(\cdot, \boldsymbol{\theta})$. Consider hypotheses

$$H_0: \ \boldsymbol{\theta} \in \Theta_0 \quad \text{vs} \quad H_1: \ \boldsymbol{\theta} \in \Theta - \Theta_0.$$

The likelihood ratio test will reject $H_0$ for the large values of the statistic

$$\begin{aligned} LR = LR(\mathbf{X}) \quad &\equiv \quad \frac{\sup_{\boldsymbol{\theta}\in\Theta} f(\mathbf{X}, \boldsymbol{\theta})}{\sup_{\boldsymbol{\theta}\in\Theta_0} f(\mathbf{X}, \boldsymbol{\theta})} \\ &= \quad f(\mathbf{X}, \widehat{\boldsymbol{\theta}})/f(\mathbf{X}, \widetilde{\boldsymbol{\theta}}), \end{aligned}$$

where $\widehat{\boldsymbol{\theta}}$ the (unconstrained) MLE, and $\widetilde{\boldsymbol{\theta}}$ is the constrained MLE under hypothesis $H_0$.

**R**emark. (i) It is easy to see that $LR \geq 1$.

(ii) The exact sampling distributions of $LR$ are usually unknown, except in a few special cases.

**E**xample 7. **(**One-sample $t$-test)
Let $\mathbf{X} = (X_1, \cdots, X_n)^\tau$ be a random sample from $N(\mu, \sigma^2)$. We are interested in testing hypotheses

$$H_0: \ \mu = \mu_0 \quad \text{against} \quad H_1: \ \mu \neq \mu_0,$$

where $\mu_0$ is given, and $\sigma^2$ is unknown and is a nuisance parameter. Now both $H_0$ and $H_1$ are composite. The likelihood function is

$$L(\mu, \sigma^2) = C\sigma^{-n} \exp\big\{ -\frac{1}{2}\sigma^2 \sum_{j=1}^n (X_i - \mu)^2 \big\}.$$

The unconstrained MLEs are

$$\widehat{\mu} = \bar{X}, \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^n (X_i - \bar{X})^2,$$

and the constrained MLE is

$$\widetilde{\sigma}^2 = \frac{1}{n}\sum_{j=1}^n (X_i - \mu_0)^2.$$

The LR-ratio statistic is then

$$LR = \frac{L(\widehat{\mu}, \widehat{\sigma}^2)}{L(\mu_0, \widetilde{\sigma}^2)} = (\widetilde{\sigma}^2/\widehat{\sigma}^2)^{n/2}.$$

Since

$$n\widetilde{\sigma}^2 = n\widehat{\sigma}^2 + n(\bar{X} - \mu_0)^2,$$

it holds that $\widetilde{\sigma}^2/\widehat{\sigma}^2 = 1 + T^2/(n-1)$, where

$$T = \sqrt{n}(\bar{X} - \mu_0)/\big\{\frac{1}{n-1}\sum_{j=1}^n (X_i - \bar{X})^2\big\}^{1/2}.$$

Note that $T \sim t_{n-1}$ under $H_0$. The LRT will reject $H_0$ iff $|T| > t_{n-1,\alpha/2}$, where $t_{k,\alpha}$ is the upper $\alpha$-point of the $t$-distribution with $k$ degrees of freedom.

**A**symptotic Distribution of Likelihood ratio test statistic
Let $\mathbf{X} = (X_1, \cdots, X_n)^\tau$, and assume certain regularity conditions. Then as $n \to \infty$, the distribution

of $2\log(LR)$ under $H_0$ converges to the $\chi^2$-distribution with $d - d_0$ degrees of freedom, where $d$ is the 'dimension' of $\Theta$ and $d_0$ is the 'dimension' of $\Theta_0$.

To make the computation of 'dimension' easy, **r**eparametrisation is often adopted. Suppose that the parameter $\theta$ may be written in two parts

$$\boldsymbol{\theta} = (\boldsymbol{\psi}, \boldsymbol{\lambda})$$

where $\boldsymbol{\psi}$ is $k \times 1$ parameter of interest, and $\boldsymbol{\lambda}$ is of little interest and is called *nuisance parameters*. The hypotheses to be tested may be expressed as

$$H_0: \ \boldsymbol{\psi} = \boldsymbol{\psi}_0 \quad \text{vs} \quad H_1: \ \boldsymbol{\psi} \neq \boldsymbol{\psi}_0.$$

Now the LR-statistic is of the form

$$LR = \frac{L(\widehat{\boldsymbol{\psi}}, \widehat{\boldsymbol{\lambda}}; \mathbf{X})}{L(\boldsymbol{\psi}_0, \widetilde{\boldsymbol{\lambda}}; \mathbf{X})},$$

where $(\widehat{\boldsymbol{\psi}}, \widehat{\boldsymbol{\lambda}})$ is unconstrained MLE while $\widetilde{\boldsymbol{\lambda}}$ is the constrained MLE of $\boldsymbol{\lambda}$ subject to $\boldsymbol{\psi} = \boldsymbol{\psi}_0$. Then as $n \to \infty$,

$$2\log(LR) \xrightarrow{D} \chi^2_k \quad \text{under } H_0.$$

**E**xample 8. Let $X_1, \cdots, X_n$ be independent, and $X_j \sim N(\mu_j, 1)$. Consider the null hypothesis

$$H_0: \ \mu_1 = \cdots = \mu_n.$$

The likelihood function is

$$L(\mu_1, \cdots, \mu_n) = C \exp\Big\{ -\frac{1}{2}\sum_{j=1}^{n}(X_j - \mu_j)^2 \Big\},$$

where $C > 0$ is a constant independent of $\mu_j$. Then the unconstrained MLE are $\widehat{\mu}_j = X_j$ and the constrained MLE is $\widetilde{\mu} = \bar{X}$. Hence

$$\begin{aligned} LR &= \frac{L(\widehat{\mu}_1, \cdots, \widehat{\mu}_n)}{L(\widetilde{\mu}, \cdots, \widetilde{\mu})} \\ &= \exp\Big\{ \frac{1}{2}\sum_{j=1}^{n}(X_j - \bar{X})^2 \Big\}. \end{aligned}$$

Hence

$$2\log(LR) = \sum_{j=1}^{n}(X_j - \bar{X})^2 \sim \chi^2_{n-1} \quad \text{under } H_0,$$

which is true for any finite $n$ as well.

How to *calculate the degree of freedom*?

Since

$$d = n, \qquad d_0 = 1,$$

the d.f. is $d - d_0 = n - 1$.

Alternatively we may adopt the following
reparametrisation:

$$\mu_j = \mu_1 + \psi_j \quad \text{for } 2 \leq j \leq n.$$

Then the null hypothesis can be expressed as

$$H_0: \ \psi_2 = \cdots = \psi_n = 0.$$

Therefore $\boldsymbol{\psi} = (\psi_2, \cdots, \psi_n)^\tau$ has $n-1$ component, i.e. $k = n-1$.