7. Simple hypotheses

A simple hypothesis $H$ specifies $f$ completely (eg $H_0: \theta = 1/2$ in (a)). Otherwise $H$ is a composite hypothesis (eg $H_1: \theta \neq 1/2$ in (b)).

For testing $H_0$ against an alternative hypothesis $H_1$, a test procedure has to partition $\mathcal{X}^n$ into two disjoint and exhaustive regions $C$ and $\bar{C}$, such that if $x \in C$ then $H_0$ is rejected and if $x \in \bar{C}$ then $H_0$ is not rejected.

The critical region (or rejection region) $C$ defines the test. When performing a test we may (i) arrive at a correct conclusion, or (ii) make one of two types of error:

(a) we may reject $H_0$ when $H_0$ is true (a Type I error).
(b) we may not reject $H_0$ when $H_0$ is false (a Type II error).

NB: When Neyman and Pearson developed the theory in the 1930s, they spoke of ‘accepting’ $H_0$. Now we generally refer to ‘not rejecting $H_0$’.

Testing a simple hypothesis against a simple alternative

When $H_0$ and $H_1$ are both simple, let

$$\alpha = P(\text{Type I error}) = P(x \in C | H_0 \text{ is true})$$
$$\beta = P(\text{Type II error}) = P(x \notin C | H_1 \text{ is true})$$

We define the size of the test to be $\alpha$.

$1 - \beta$ is also known as the power of the test to detect $H_1$. Ideally we would like $\alpha = \beta = 0$, but typically it is not possible to find a test that makes both $\alpha$ and $\beta$ arbitrarily small.

**Definition 7.1**

- The likelihood of a simple hypothesis $H: \theta = \theta^*$ given data $x$ is $L_x(H) = f(x | \theta = \theta^*)$.
- The likelihood ratio of two simple hypotheses $H_0, H_1$, given data $x$, is $\Lambda_x(H_0, H_1) = L_x(H_1)/L_x(H_0)$.
- A likelihood ratio test (LR test) is one where the critical region $C$ is of the form $C = \{x : \Lambda_x(H_0, H_1) > k\}$ for some $k$. □

Let $X_1, \ldots, X_n$ be iid, each taking values in $\mathcal{X}$, each with unknown pdf/pmf $f$, and suppose that we have two hypotheses, $H_0$ and $H_1$, about $f$.

On the basis of data $X = x$, we make a choice between the two hypotheses.

**Examples**

(a) A coin has $P(\text{Heads}) = \theta$, and is thrown independently $n$ times. We could have $H_0 : \theta = 1/2$ versus $H_1 : \theta = 3/4$.
(b) As in (a), with $H_0 : \theta = 1/2$ as before, but with $H_1 : \theta \neq 1/2$.
(c) Suppose $X_1, \ldots, X_n$ are iid discrete rv’s. We could have $H_0$ : the distribution is Poisson with unknown mean, and $H_1$ : the distribution is not Poisson. This is a goodness-of-fit test.
(d) General parametric case: $X_1, \ldots, X_n$ are iid with density $f(x | \theta)$, with $H_0: \theta \in \Theta_0$ and $H_1: \theta \in \Theta_1$ where $\Theta_0 \cap \Theta_1 = \emptyset$ (we may or may not have $\Theta_0 \cup \Theta_1 = \Theta$).
(e) We could have $H_0: f = f_0$ and $H_1: f = f_1$ where $f_0$ and $f_1$ are densities that are completely specified but do not come from the same parametric family.
### Proof

The given $C$ specifies a likelihood ratio test with size $\alpha$.

Let $\beta = \mathbb{P}(X \not\in C \mid f_1) = \int_{\bar{C}} f_1(x) \, dx$.

Let $C^*$ be the critical region of any other test with size less than or equal to $\alpha$.

Let $\alpha^* = \mathbb{P}(X \in C^* \mid f_0)$, $\beta^* = \mathbb{P}(X \not\in C^* \mid f_1)$.

We want to show $\beta \leq \beta^*$.

We know $\alpha^* \leq \alpha$, i.e. $\int_{\bar{C}} f_0(x) \, dx \leq \int_{\bar{C}} f_0(x) \, dx$.

Also, on $C$ we have $f_1(x) > k f_0(x)$, while on $\bar{C}$ we have $f_1(x) \leq k f_0(x)$.

Thus

$$\int_{C} f_0(x) \, dx \leq k \int_{C} f_0(x) \, dx, \quad \int_{C} f_0(x) \, dx \leq k \int_{C} f_0(x) \, dx. \quad (1)$$

Hence

\[
\beta - \beta^* = \int_{C} f_1(x) \, dx - \int_{C^*} f_1(x) \, dx = \int_{\bar{C}} f_1(x) \, dx + \int_{C \cap C^*} f_1(x) \, dx - \int_{C^*} f_1(x) \, dx \\
\leq k \int_{\bar{C}} f_0(x) \, dx - k \int_{C \cap C^*} f_0(x) \, dx = k \left( \int_{\bar{C}} f_0(x) \, dx + \int_{C \cap C^*} f_0(x) \, dx \right) - k \left( \int_{C \cap C^*} f_0(x) \, dx + \int_{C^*} f_0(x) \, dx \right) \\
= k (\alpha^* - \alpha) \leq 0.
\]

\[\square\]

### Example 7.3

Suppose that $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2_0)$, where $\sigma^2_0$ is known. We want to find the best size $\alpha$ test of $H_0: \mu = \mu_0$ against $H_1: \mu = \mu_1$, where $\mu_0$ and $\mu_1$ are known fixed values with $\mu_1 > \mu_0$.

\[
\Lambda_\alpha(H_0; H_1) = \frac{(2\pi\sigma^2_0)^{-n/2} \exp \left(-\frac{1}{2\sigma^2_0} \sum (x_i - \mu_1)^2 \right)}{(2\pi\sigma^2_0)^{-n/2} \exp \left(-\frac{1}{2\sigma^2_0} \sum (x_i - \mu_0)^2 \right)} = \exp \left( \frac{(\mu_1 - \mu_0) n \bar{x} + n(\mu_0^2 - \mu_1^2)}{2\sigma^2_0} \right) \quad (\text{check}).
\]

This is an increasing function of $\bar{x}$, so for any $k$,

\[\Lambda_\alpha > k \Leftrightarrow \bar{x} > c \text{ for some } c.
\]

Hence we reject $H_0$ if $\bar{x} > c$ where $c$ is chosen such that $\mathbb{P}(\bar{X} > c \mid H_0) = \alpha$.

Under $H_0$, $\bar{x} \sim N(\mu_0, \sigma^2_0/n)$, so $Z = \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 \sim N(0,1)$.

Since $\bar{x} > c \Leftrightarrow Z > c'$ for some $c'$, the size $\alpha$ test rejects $H_0$ if $Z = \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 > z_\alpha$. 

- We assume continuous densities to ensure that a LR test of exactly size $\alpha$ exists.
- The Neyman–Pearson Lemma shows that $\alpha$ and $\beta$ cannot both be arbitrarily small.
- It says that the most powerful test (ie the one with the smallest Type II error probability), among tests with size smaller than or equal to $\alpha$, is the size $\alpha$ likelihood ratio test.
- Thus we should fix $\mathbb{P}(\text{Type I error})$ at some level $\alpha$ and then use the Neyman–Pearson Lemma to find the best test.
- Here the hypotheses are not treated symmetrically; $H_0$ has precedence over $H_1$ and a Type I error is treated as more serious than a Type II error.
- $H_0$ is called the null hypothesis and $H_1$ is called the alternative hypothesis.
- The null hypothesis is a conservative hypothesis, ie one of “no change,” “no bias,” “no association,” and is only rejected if we have clear evidence against it.
- $H_1$ represents the kind of departure from $H_0$ that is of interest to us.
7. Simple hypotheses

7.2. Testing a simple hypothesis against a simple alternative

Suppose $\mu_0 = 5, \mu_1 = 6, \sigma_0 = 1, \alpha = 0.05, n = 4$ and $x = (5.1, 5.5, 4.9, 5.3)$, so that $\bar{x} = 5.2$.

- From tables, $z_{0.05} = 1.645$.
- We have $z = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma_0} = 0.4$ and this is less than 1.645, so $x$ is not in the rejection region.
- We do not reject $H_0$ at the 5%- level; the data are consistent with $H_0$.
- This does not mean that $H_0$ is 'true', just that it cannot be ruled out.
- This is called a $z$-test.

P-values

- In this example, LR tests reject $H_0$ if $z > k$ for some constant $k$.
- The size of such a test is $\alpha = P(Z > k \mid H_0) = 1 - \Phi(k)$, and is decreasing as $k$ increases.
- Our observed value $z$ will be in the rejection region $\iff z > k \iff \alpha > p^* = P(Z > z \mid H_0)$.
- The quantity $p^*$ is called the $p$-value of our observed data $x$.
- For Example 7.3, $z = 0.4$ and so $p^* = 1 - \Phi(0.4) = 0.3446$.
- In general, the $p$-value is sometimes called the 'observed significance level' of $x$ and is the probability under $H_0$ of seeing data that are 'more extreme' than our observed data $x$.
- Extreme observations are viewed as providing evidence against $H_0$.
- The $p$-value has a Uniform(0,1) pdf under the null hypothesis. To see this for a $z$-test, note that
  
  $P(p^* < p \mid H_0) = P([1 - \Phi(Z)] < p \mid H_0) = P(Z > \Phi^{-1}(1 - p) \mid H_0)$
  
  $= 1 - \Phi(\Phi^{-1}(1 - p)) = 1 - (1 - p) = p.$