Finite sample testing theory

Bent Nielsen

University of Oxford

6 November 2023

Comments welcome — mistakes present with probability 1

Slightly revised version of slides written by Anders Kock 2022

Introduction to testing

- So far we have studied estimation of θ in the context of the experiment

$$\mathcal{E} = \left(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\right).$$

- We shall now turn towards testing hypotheses about θ.

The presentation is primarily based on Section 2.2 in Liese and Miescke (2008), Section 4.1–4.3 in Pfanzagl (1994) and Sections 3.1–3.2 as well as 3.4 in Lehmann and Romano (2005). The notation is largely as in Liese and Miescke (2008) and some examples are taken form Witting (1985), pp 110–113.

Plan

- Testing framework: We begin by introducing/reviewing the basic concepts of testing.
- We establish the Neyman-Pearson lemma, which characterizes most powerful tests of a prespecified level when testing a simple null against a simple alternative.
- Typically, the most powerful test (of a given level) depends on the alternative. However, in one-parameter families with monotone likelihood ratios, uniformly most powerful tests for one-sided hypotheses against one-sided alternatives exist. This is the content of the Karlin-Rubin theorem.
    - One-parameter exponential families often have monotone likelihood ratios.

The framework

- In the context of the experiment

$$\mathcal{E} = \left(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\right),$$

  one often wishes to establish whether θ belongs to one of the two disjoint subsets Θ_0 or Θ_1 of Θ.
- Thus, one formulates the hypotheses

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

- H_0 is called the *null hypothesis* and H_1 is called the *alternative hypothesis*.
- One also frequently writes H_A for H_1; the A standing for *alternative*.

Example

In the context of $\mathsf{N}(\mu, \sigma_0^2)$ with $\sigma_0^2 > 0$ known and $\mu \in \Theta = \mathbb{R}$, one may wish to test

$$H_0 : \mu \in (-\infty, 0] \quad \text{against} \quad H_1 : \mu \in (0, \infty).$$

This corresponds to $\Theta_0 = (-\infty, 0]$ and $\Theta_1 = (0, \infty)$.

Test

- The typical tool for making a decision between H_0 and H_1 is a *test*.
- As always, we equip subsets of the real line with the Borel σ-algebra.

Definition 1 (Test)

A function $\varphi : \mathcal{X} \to [0, 1]$ that is $\mathcal{A}$-$\mathcal{B}([0, 1])$-measurable is called a *test*.

Comments on test

- If a test only takes the values 0 and 1, that is if $\varphi(\mathcal{X}) \subseteq \{0,1\}$, it is called non-randomized.
- Otherwise, the test is called randomized.
- Reason: Upon observing $x \in \mathcal{X}$, we interpret $\varphi(x)$ as the probability of rejecting H_0.
- Most tests you have encountered are non-randomized.
- A non-randomized test can be written as $\varphi(x) = \mathbb{1}_C(x)$ for a $C \in \mathcal{A}$.
- Observe that we reject if and only if $x \in C$.
- Hence, C is called the *rejection region* or the *critical region*.
- The complement of C is called the *acceptance region*.

Test statistics

- Often a test is based on a *test statistic* $T : \mathcal{X} \to \mathbb{R}$ that is $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-measurable.
- And we reject if $T(x)$ exceeds a *critical value* $c \in \mathbb{R}$ (chosen to ensure, e.g., a given size of the test)
- This corresponds to the non-randomized test

$$\varphi_{T,c}(x) = \mathbb{1}_{\{u \in \mathcal{X} : T(u) > c\}}(x).$$

- That is the critical region is $C = \{x \in \mathcal{X} : T(x) > c\}$.
- Observe that $C = T^{-1}(c, \infty) \in \mathcal{A}$ by the $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-measurability of T.
- Thus, the classical tests based on a test statistic and a critical value are non-randomized.

Example

- Assume we observe n independent observation from $N(\mu, \sigma_0^2)$ with $\sigma_0^2 > 0$ known and $\mu \in \Theta = \mathbb{R}$.
- In testing

$$H_0 : \mu \in (-\infty, 0] \quad \text{against} \quad H_1 : \mu \in (0, \infty),$$

The test $T(x) = \frac{\bar{x}}{\sigma_0/\sqrt{n}}$ with critical value 1.645 has rejection probability under the null of 5%
since $N(0, \sigma_0^2)^n \circ T^{-1} = N(0, 1)$.
- Thus, the critical region is $C = \left\{x \in \mathbb{R}^n : \frac{\bar{x}}{\sigma_0/\sqrt{n}} > 1.645\right\}$ and the non-randomized test is $\mathbb{1}_C(x)$.

The role of randomized tests

- Although most tests used in practice are non-randomized, randomized test play a crucial role in constructing most powerful tests of a prespecified level.
- We shall see (and show in the exercises) that this is because randomization allows us to exhaust the level constraint.
- This is important, in particular, for data from discrete distributions.

Power function

Definition 2

The function $\beta : \Theta \to [0, 1]$ defined via $\theta \mapsto E_\theta\varphi$ is called the *power function* of the test φ.

- We want tests for which $\beta(\theta)$ is "as large as possible" on Θ_1 and "as small as possible" on Θ_0.
- We call $\beta(\theta)$ the power of φ against $\theta \in \Theta_1$.
- Deciding in favor of H_1 when H_0 is true is called a *Type 1 error* (error of the first kind). For $\theta \in \Theta_0$ the probability of this is $\beta(\theta)$.
- Deciding in favor of H_0 when H_1 is true is called a *Type 2 error* (error of the second kind). For $\theta \in \Theta_1$ the probability of this is $1 - \beta(\theta)$.

Binary experiments and simple hypotheses

- In the simplest testing problem there are only two distributions P_0 and P_1. Thus, the *binary* experiment is

$$\mathcal{E} = \left(\mathcal{X}, \mathcal{A}, \{P_0, P_1\}\right),$$

and one tests

$$H_0 : P_0 \quad \text{against} \quad H_1 : P_1$$

- The null and alternative are here called *simple* since they consist of one element only.
- Otherwise, they are called *composite*.

General non-existence of a best test

- Denoting by E_0 and E_1 the expectations corresponding to P_0 and P_1, respectively, an ideal test φ would satisfy

$$E_0\varphi = 0 \quad \text{and} \quad E_1\varphi = 1.$$

- Unfortunately, such an ideal test only exists in trivial situations.
- To establish this, we introduce the concept of measures being singular.

Definition 3 (Singular measures)

Two measures P_0 and P_1 on $(\mathcal{X}, \mathcal{A})$ are singular if there exists an $A \in \mathcal{A}$ such that $P_0(A) = 0$ while $P_1(A^c) = 0$ (or if P_1 is a probability measure $P_1(A) = 1$).

- Singular measures are, in a sense, the opposite of mutually absolutely continuous measures.
- There exists a P_0-null set on which P_1 has all its mass and vice versa.

Theorem 4 (General non-existence of best test)

Consider testing $H_0 : P_0$ against $H_1 : P_1$ in the binary experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_0, P_1\}\big)$. There exists a test φ with

$$E_0\varphi = 0 \quad \textit{and} \quad E_1\varphi = 1.$$

if and only if P_0 and P_1 are singular.

Proof

If P_0 and P_1 are singular there exists an $A \in \mathcal{A}$ such that $P_0(A) = 0$ and $P_1(A) = 1$. Thus, the test $\varphi = \mathbb{1}_A$ satisfies $E_0\varphi = P_0(A) = 0$ and $E_1\varphi = P_1(A) = 1$ as desired.

Conversely, assume that there exists a test φ such that $E_0\varphi = 0$ and $E_1\varphi = 1$. Set $A = \{\varphi = 1\}$. Then, $P_0(A) = P_0(\varphi = 1) = 0$ since otherwise $E_0\varphi > 0$. Furthermore, $P_1(A) = P_1(\varphi = 1) = 1$ since otherwise $E_1\varphi < 1$.

$\square$

Comments

- Clearly, if a best test generally does not exist when testing against a simple alternative, no best test (a test having power one against all elements in Θ_1) generally exists when testing against a composite alternative.
  [Simply observe that P_1 would be one of many alternatives so we can't have power one against all of these if we don't have power one against P_1.]
- Sometimes, the above theorem can be used constructively.
- Consider testing

$$H_0 : \mathrm{N}(\mu_0, \sigma_0^2) \quad \text{against} \quad H_1 : \mathrm{Poi}(\lambda_1),$$

  for fixed $\mu_0 \in \mathbb{R}, \sigma_0^2 > 0$ and $\lambda_1 > 0$.
- These two measures are singular since $\mathrm{N}(\mu_0, \sigma_0^2)(\mathbb{Z}) = 0$ while $\mathrm{Poi}(\lambda_1)(\mathbb{Z}) = 1$.
- Hence, we can use the test $\varphi = \mathbb{1}_{\mathbb{Z}}$, which clearly has a rejection probability under H_0 of 0 and 1 under H_1.

Example: Absolutely continuous measures

- If, on the other hand, we test

$$H_0 : \mathrm{N}(\mu_0, \sigma_0^2) \quad \text{against} \quad H_1 : \mathrm{N}(\mu_1, \sigma_0^2),$$

with $\sigma_0^2 > 0$ known, we observe that $\mathrm{N}(\mu_0, \sigma_0^2)$ and $\mathrm{N}(\mu_1, \sigma_0^2)$ are mutually absolutely continuous for any pair $(\mu_0, \mu_1) \in \mathbb{R}^2$.

- Hence, for all $A \in \mathcal{A}$ one has

$$\mathrm{N}(\mu_0, \sigma_0^2)(A) = 0 \quad \Leftrightarrow \quad \mathrm{N}(\mu_1, \sigma_0^2)(A) = 0.$$

- Therefore, for any test φ

$$E_0\varphi = 0 \Rightarrow \mathrm{N}(\mu_0, \sigma_0^2)(\varphi \neq 0) = 0 \Rightarrow \mathrm{N}(\mu_1, \sigma_0^2)(\varphi \neq 0) = 0,$$

implying that $E_1\varphi = 0$.

- Thus, any test of with rejection probability 0 under the null, also has rejection probability 0 under the alternative.
- This is the case, *no matter how far* μ_0 and μ_1 are apart.

- Let us also recall that exponential families are mutually absolutely continuous.
- Hence, we can not have $E_\theta \varphi = 0$ for $\theta \in \Theta_0$ without having $E_\theta \varphi = 0$ for $\theta \in \Theta_1$.
- Thus, requiring $E_\theta \varphi = 0$ for $\theta \in \Theta_0$ is too much.

Restricting the class of tests

- Since there is generally no optimal test, we restrict the class of tests in order to build a meaningful theory of testing.
- One approach is to maximize the power of the test under the constraint that the rejection probability under the null does not exceed a prespecified value.
- This approach amounts to a strong desire to avoid Type 1 errors since one often chooses the allowed rejection probability under H_0 to be rather low.
- ...and only given this constraint minimizes the Type 2 error probability (i.e. maximizes power).

Definition 5 (Size and level)

For any test $\varphi : \mathcal{X} \to [0,1]$ its size is defined to be

$$\alpha(\varphi) := \sup_{\theta \in \Theta_0} E_\theta \varphi.$$

For $\alpha \in (0,1)$ a test φ is called a level α test for H_0 if $\alpha(\varphi) \leq \alpha$. Especially we say that a test φ attains the level α if $\alpha(\varphi) = \alpha$.

- It makes good sense not require $E_\theta \varphi = 0$ for any $\theta \in \Theta_0$ since we have seen that this implies power equal to zero against absolutely continuous alternatives.
- Thus, it is “natural” to allow for strictly positive size.
- This is the *Neyman-Pearson* testing framework.

Best tests in subclasses and UMP tests

Definition 6

Let $\mathbb{T}_0$ be a subclass of tests. A test ψ is called a uniformly best test in $\mathbb{T}_0$ for H_0 versus H_1 if $\psi \in \mathbb{T}_0$ and for every $\varphi \in \mathbb{T}_0$ it holds

$$E_\theta \varphi \leq E_\theta \psi \qquad \text{for all } \theta \in \Theta_1.$$

Especially, if $\mathbb{T}_0 = \{\varphi : \alpha(\varphi) \leq \alpha\}$, for some $\alpha \in (0,1)$, is the set of level α tests, then ψ is called a uniformly best level α test.

- Uniformly best is also called *uniformly most powerful* (UMP).
- UMP tests exist under special circumstances only as the identity of the most powerful test typically depends on $\theta \in \Theta_1$.

We shall see in the exercises that unless a UMP test ψ at level $\alpha \in (0,1)$ has power 1 throughout H_1, that is $E_\theta \psi = 1$ for all $\theta \in \Theta_1$, it always exhausts the level constraint, that is $\alpha(\psi) = \alpha$.

Towards the Neyman-Pearson Lemma

- The Neyman-Pearson Lemma characterizes the optimal tests in testing $H_0 : P_0$ against $H_1 : P_1$ in a binary experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_0, P_1\}\big)$. [Simple null and simple alternative.]
- However, we shall see that its consequence reach further:
  - It gives an upper bound on the power that any test of level $\alpha \in (0, 1)$ can achieve.
  - It is important in constructing UMP tests in one-parameter one-sided testing problems.

Some heuristics

- To build some intuition, let $\mathcal{X}$ be a finite set and restrict attention to nonrandomized tests.
- Thus, a test amounts to choosing a critical region $C \subseteq \mathcal{X}$ for which we reject H_0.
- For a level α test, we wish to choose C such that

$$P_1(C) = \sum_{x \in C} P_1(x) \text{ is maximized}$$

  subject to $P_0(C) = \sum_{x \in C} P_0(x) \leq \alpha$.
- Hence, we wish to choose the elements x of C such that $P_0(x)$ is small while $P_1(x)$ is large.
- Certainly, any x with $P_0(x) = 0$ should be included in C as it "costs nothing" in terms of size, but contributes to power if $P_1(x) > 0$.

Given the above heuristics, it is not surprising that the most powerful test for P_0 versus P_1 chooses

$$C = \left\{ x \in \mathcal{X} : \frac{P_1(x)}{P_0(x)} \geq c \right\},$$

where $c > 0$ is determined by the condition

$$P_0(C) = \sum_{x \in C} P_0(x) \leq \alpha.$$

[where $\frac{a}{0}$ is interpreted as ∞ for $a > 0$ and 0 for $a = 0$.]

Likelihood ratios — partial recap

Recall that if μ and ν are two σ-finite measures on $(\mathcal{X}, \mathcal{A})$, such that ν is absolutely continuous with respect to μ, that is for every $A \in \mathcal{A}$ one has that $\mu(A) = 0$ implies $\nu(A) = 0$, then there exists a $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-measurable $f : \mathcal{X} \to \mathbb{R}$ with $f \geq 0$ such that

$$\nu(B) = \int_B f(x)\mu(dx) \qquad \text{for every } B \in \mathcal{A}. \tag{1}$$

Any f with the above property in (1) is called a *density*, the *Radon-Nikodym derivative* or the *likelihood ratio* of ν with respect to μ.

One often writes $f = \frac{d\nu}{d\mu}$.

We write $\nu \ll \mu$ to indicate that ν is absolutely continuous with respect to μ.

Example

- Let $\mathcal{X} = \mathbb{R}$ and $\mathcal{A} = \mathcal{B}(\mathbb{R})$.
- Let $\nu = \mathrm{N}(\eta, \sigma^2)$ and $\mu = \lambda_1$, the Lebesgue measure on $\mathcal{B}(\mathbb{R})$.
- Then ν is absolutely continuous with respect to λ_1 and

$$\frac{d\nu}{d\lambda_1}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\eta)^2}{2\sigma^2}},$$

that is, the usual *density* of the normal distribution.

- Most commonly used distributions possess a density (with respect to Lebesgue measure) or a probability mass functions (with respect to counting measure).
- Indeed, we often take these as the definitions of the distributions.
- However, it is useful to be able to find the likelihood ratio of, say, one normal distribution with respect to another, say $\mathsf{N}(\mu_1, \sigma^2)$ with respect to $\mathsf{N}(\mu_0, \sigma^2)$.
- We have the following general rule that you have established in the exercises:
  If for σ-finite measures $\nu \ll \mu \ll \rho$, then

$$\frac{d\nu}{d\mu} = \left(\frac{d\nu}{d\rho}\Big/\frac{d\mu}{d\rho}\right) \times \mathbb{1}_{\left\{\frac{d\mu}{d\rho}>0\right\}} \qquad \mu\text{-a.e.} \tag{2}$$

Example: Likelihood ratio of normal distributions

Thus, using $\nu = P_1,\ \mu = P_0$ and $\rho = \lambda_1$ as well as using that densities of normal distributions are everywhere strictly positive, the above rule gives

$$\frac{dP_1}{dP_0} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma^2}} \Big/ \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_0)^2}{2\sigma^2}} = e^{\frac{(x-\mu_0)^2}{2\sigma^2} - \frac{(x-\mu_1)^2}{2\sigma^2}}$$
$$= e^{\frac{x(\mu_1-\mu_0)}{\sigma^2} - \frac{\mu_1^2-\mu_0^2}{2\sigma^2}}.$$

Uniform distribution

- Let $\mathcal{X} = \mathbb{R}$, $\mathcal{A} = \mathcal{B}(\mathbb{R})$, $P_0 = \mathrm{U}(0,1)$ and $P_1 = \mathrm{U}(0,\theta)$ for $\theta > 0$.
- Thus,

$$p_0(x) := \frac{dP_0}{d\lambda_1}(x) = \mathbb{1}_{[0,1]}(x) \quad \text{and} \quad p_1(x) := \frac{dP_1}{d\lambda_1}(x) = \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(x)$$

- If $\theta > 1$, then we do *not* have that $P_1 \ll P_0$ since $P_0([1,\theta]) = 0$, but $P_1([1,\theta]) = \frac{\theta-1}{\theta} > 0$.
- Thus, there exists no density of P_1 wrt. P_0.
- This is annoying since a general likelihood ratio testing theory should encompass this simple situation, cf. next slide.

Heuristics/motivation

- Note, however, that if we observe an $x \in (1, \theta]$, then we are certain that the distribution is P_1 since $P_0((1, \theta]) = 0$ but $P_1((1, \theta]) > 0$.
- Since the likelihood ratio reflects "how likely P_1 is compared to P_0" and since we decide in favor of P_1 when p_1/p_0 is large, it is intuitively sensible to set it equal to ∞ when $p_1 > 0$ but $p_0 = 0$.
- Note $(p_0 > 0)$ is short hand for $\{x : p_0(x) > 0\}$.
- We often use the notion of a likelihood ratio test statistic for the ratio of two maxima of a likelihood, maximized under a restricted model (a hypothesis) and under an unrestricted model. We call that a maximum likelihood ratio test statistic if we want to distinguish from the present likelihood ratios.

Likelihood ratio

- Note that the following definition does *not* require $P_1 \ll P_0$.
- Nothing is lost in imposing $P_0 \ll \mu$ and $P_1 \ll \mu$ for some μ since we can always choose $\mu = P_0 + P_1$.
- Thus, the definition applies to *all* pairs P_0 and P_1 of probability measures.

Definition 7 (Likelihood ratio)

Let P_0 and P_1 be probability measures on $(\mathcal{X}, \mathcal{A})$ that are absolutely continuous with respect to a σ-finite measure μ, with densities p_0 and p_1, respectively. Define $L_{0,1} : \mathcal{X} \to [0, \infty]$ via

$$L_{0,1}(x) := \frac{p_1(x)}{p_0(x)} \mathbb{1}_{\{p_0 > 0\}}(x) + \infty \cdot \mathbb{1}_{\{p_0 = 0, p_1 > 0\}}(x).$$

Then $L_{0,1}$ is the a likelihood ratio of P_1 with respect to P_0.

The convention $\infty \cdot 0 = 0$ is used throughout.

We check that this Definition 7 satisfies the conditions for a Radon-Nikodym/likelihood ratio (page 25) when $P_1 \ll P_0$:

$$P_0(p_0 = 0) = \int \mathbb{1}_{\{p_0=0\}} P_0(dx) = \int \mathbb{1}_{\{p_0=0\}} p_0(x)\mu(dx) = 0,$$

and so also $P_0(p_0 = 0, p_1 > 0) \leq P_0(p_0 = 0) = 0$.

Hence, for any $A \in \mathcal{A}$, if $P_1 \ll P_0$

$$\begin{aligned}
&\int_A L_{0,1}(x) P_0(dx) \\
=&\int_A \frac{p_1(x)}{p_0(x)} \mathbb{1}_{\{p_0>0\}}(x) P_0(dx) + \infty \cdot P_0(A \cap \{p_0 = 0, p_1 > 0\}) \\
=&P_1(A \cap \{p_0 > 0\}) \\
=&P_1(A),
\end{aligned}$$

since $P_1(A \cap \{p_0 = 0\}) \leq P_1(p_0 = 0) = 0$ because $P_0(p_0 = 0) = 0$ and $P_1 \ll P_0$ by assumption.

Example: Normal distribution

- Consider $P_0 = \mathrm{N}(\mu_0, \sigma^2)$ and $P_1 = \mathrm{N}(\mu_1, \sigma^2)$ with $\sigma^2 > 0$.
- The corresponding densities p_0 and p_1 (with respect to the Lebesgue measure λ_1) are everywhere strictly positive.
- Hence,

$$L_{0,1}(x) = \frac{p_1(x)}{p_0(x)} \mathbb{1}_{\{p_0>0\}}(x) + \infty \cdot \mathbb{1}_{\{p_0=0, p_1>0\}}(x) = \frac{p_1(x)}{p_0(x)},$$

as before.

Example: Uniform distribution

- With the generalized definition of a likelihood ratio, likelihood ratios between *all* pairs of probability measures exist.
- In particular, we can now handle uniform distributions.

Let $P_0 = \mathrm{U}(0,1)$ and $P_1 = \mathrm{U}(0,\theta)$ for $\theta > 1$ and recall that

$$p_0(x) = \frac{dP_0}{d\lambda_1}(x) = \mathbb{1}_{[0,1]}(x) \quad \text{and} \quad p_1(x) = \frac{dP_1}{d\lambda_1}(x) = \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(x).$$

Thus,

$$\begin{aligned} L_{0,1}(x) &= \frac{p_1(x)}{p_0(x)}\mathbb{1}_{\{p_0>0\}}(x) + \infty \cdot \mathbb{1}_{\{p_0=0,p_1>0\}}(x) \\ &= \frac{1}{\theta}\mathbb{1}_{[0,1]}(x) + \infty \cdot \mathbb{1}_{(1,\theta]}(x). \end{aligned}$$

This is sensible since we know that if we observe $x \in (1,\theta]$, then the observation is from P_1.

Likelihood ratio tests

Definition 8 (Likelihood ratio tests)

In the binary experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_0, P_1\}\big)$ we call φ_c a *likelihood ratio* (LR) test with critical value $c \in [0, \infty)$ for $H_0 : P_0$ against $H_1 : P_1$ if

$$\begin{aligned} \varphi_c &= 1 \quad \text{on } \{p_1 > cp_0\} \\ \varphi_c &= 0 \quad \text{on } \{p_1 < cp_0\}\,. \end{aligned}$$

- Thus, a likelihood ratio tests rejects when $p_1 > cp_0$.
- While it accepts when $p_1 < cp_0$.
- Note that it is unspecified on $\{p_1 = cp_0\}$.
- On that set we may randomize to ensure that we can get size exactly α for a desired $\alpha \in (0, 1)$.

We observe that (cf. exercises)

$$P_i\left(L_{0,1} > c\right) = P_i\left(p_1 > cp_0\right) \qquad \text{for } i = 0,1$$
$$P_i\left(L_{0,1} < c\right) = P_i\left(p_1 < cp_0\right) \qquad \text{for } i = 0,1$$

and thus

$$P_i\left(L_{0,1} = c\right) = P_i\left(p_1 = cp_0\right) \qquad \text{for } i = 0,1$$

Hence, a test φ_c' with the same size and power as φ_c can be obtained by defining

$$\varphi_c' = 1 \quad \text{on } \{L_{0,1} > c\}$$
$$\varphi_c' = 0 \quad \text{on } \{L_{0,1} < c\}$$

Since $\varphi_c' = \varphi_c$ $\{P_0, P_1\}$-*a.s.*, they have the *same* power functions.

1. φ_c' better motivates the name *likelihood ratio* test as it is defined directly via the likelihood ratio $L_{0,1}$.
2. φ_c defined p_0 and p_1 is easier to use in the proof of the Neyman-Pearson Lemma (as we shall see).

φ_c' and φ_c are the "same" test as they only differ on a $\{P_0, P_1\}$-null set and hence have the same power functions.

Some observations

- In light of the above, we can equivalently say that a likelihood ratio tests rejects when $L_{0,1} > c$ and does not reject when $L_{0,1} < c$.
- Observe that since $P_0(p_0 = 0) = 0$, we have that $P_0(L_{0,1} < \infty) = 1$.
- Thus, defining F_0 as the cdf of $L_{0,1}$ under P_0, that is $F_0(t) := P_0(L_{0,1} \leq t),\ t \in \mathbb{R}$, one has $F_0^{-1}(1-\alpha) < \infty$ for all $\alpha \in (0,1)$.

With $c_{1-\alpha} := F_0^{-1}(1-\alpha)$ for $\alpha \in (0,1)$, define

$$\psi_\alpha(t) := \begin{cases} 1 & \text{if } t > c_{1-\alpha} \\ \gamma_\alpha & \text{if } t = c_{1-\alpha}, \\ 0 & \text{if } t < c_{1-\alpha} \end{cases} \qquad t \in [0,\infty]$$

where we choose γ_α so that
$\alpha = P_0(L_{0,1} > c_{1-\alpha}) + \gamma_\alpha P_0(L_{0,1} = c_{1-\alpha})$. Solve to get

$$\gamma_\alpha = \begin{cases} \frac{F_0(c_{1-\alpha})-(1-\alpha)}{P_0(L_{0,1}=c_{1-\alpha})} & \text{if } P_0(L_{0,1} = c_{1-\alpha}) > 0 \\ 0 & \text{if } P_0(L_{0,1} = c_{1-\alpha}) = 0. \end{cases} \tag{3}$$

- Observe that $\psi_\alpha(L_{0,1})$ is a Likelihood ratio test.
- It has been given a specific value on $\{L_{0,1} = c_{1-\alpha}\}$ to ensure that it has size exactly α (as we shall see).

- We stress again that assuming that P_0 and P_1 are dominated by a σ-finite measure μ is "for free" since we can always choose $\mu = P_0 + P_1$.

Theorem 9 (Neyman-Pearson Lemma)

Consider testing $H_0 : P_0$ against $H_1 : P_1$ in $\mathcal{E} = \left(\mathcal{X}, \mathcal{A}, \{P_0, P_1\}\right)$, where P_0 and P_1 are dominated by a σ-finite measure μ.

1. *For every $\alpha \in (0,1)$ there exists an LR test of size α. In particular, $E_0\psi_\alpha(L_{0,1}) = \alpha$.*
2. *Every LR test φ_c is a best test at level $\alpha(\varphi_c)$. In particular, $\psi_\alpha(L_{0,1})$ is best at level α.*
3. *For every $\alpha \in (0,1)$, every best level α test φ with $\alpha(\varphi) = \alpha$ equals $\psi_\alpha(L_{0,1})$ outside $\{L_{0,1} = c_{1-\alpha}\}$ for μ-almost all $x \in \mathcal{X}$. That is,*

$$\mu\left(\{(\psi_\alpha \circ (L_{0,1}))(x) - \varphi(x) \neq 0\} \cap \{p_1(x) - c_{1-\alpha}p_0(x) \neq 0\}\right) = 0.$$

Comments

The Neyman-Pearson Lemma has three components:

1. *Existence of a LR test of a desired size*: For any $\alpha \in (0,1)$, the lemma guarantees the existence of a LR test of that size.
2. *The LR test of size α is optimal*: That is, the LR test of size α maximizes power against P_1 in the class of level α tests.
3. If a test is optimal at level $\alpha \in (0,1)$ and exhausts the level constraint, it equals the test ψ_α on $\{p_1 \neq c_{1-\alpha} p_0\}$ for μ-almost every $x \in \mathcal{X}$.

Proof of Neyman-Pearson Lemma

First, if $P_0(L_{0,1} = c_{1-\alpha}) > 0$,

$$\begin{aligned} E_0\psi_\alpha(L_{0,1}) &= P_0(L_{0,1} > c_{1-\alpha}) + \gamma_\alpha P_0(L_{0,1} = c_{1-\alpha}) \\ &= 1 - F_0(c_{1-\alpha}) + \frac{F_0(c_{1-\alpha}) - (1-\alpha)}{P_0(L_{0,1} = c_{1-\alpha})} P_0(L_{0,1} = c_{1-\alpha}) \\ &= \alpha. \end{aligned}$$

If, on the other hand, $P_0(L_{0,1} = c_{1-\alpha}) = 0$ then

$$E_0\psi_\alpha(L_{0,1}) = 1 - F_0(c_{1-\alpha}) = 1 - F_0(F_0^{-1}(1-\alpha)) = \alpha,$$

establishing that $E_0\psi_\alpha(L_{0,1}) = \alpha$.[1]

---

[1] $P_0(L_{0,1} = c_{1-\alpha}) = 0$ is equivalent to F_0 being continuous at $c_{1-\alpha}$. Furthermore, one *always* has $F_0(c_{1-\alpha}) = F_0(F_0^{-1}(1-\alpha)) \geq 1-\alpha$ by right-continuity of F_0. If F_0 is continuous and $F_0(c_{1-\alpha}) > 1-\alpha$ this contradicts the definition of $c_{1-\alpha} = F^{-1}(1-\alpha)$ as also $F_0(F^{-1}(1-\alpha) - \delta) > 1-\alpha$ for $\delta > 0$ sufficiently small. Thus, $F_0(F_0^{-1}(1-\alpha)) = 1-\alpha$.

To see that every LR test φ_c is best at level $\alpha(\varphi_c)$, we consider any other test ϕ of level $\alpha(\varphi_c)$.
First, we argue that for any test ϕ (regardless of level)

$$(\varphi_c - \phi)(p_1 - cp_0) \geq 0 \qquad \text{for all } x \in \mathcal{X}. \tag{4}$$

Recall that tests maps into $[0,1]$.
On $p_1(x) - cp_0(x) > 0$ then $\varphi_c = 1$ and $\phi \leq 1$ and (4) holds.
On $p_1(x) - cp_0(x) = 0$ then (4) holds for any values of φ_c, ϕ.
On $p_1(x) - cp_0(x) < 0$ then $\varphi_c = 0$ and $\phi \geq 0$ and (4) holds.
Now, using (4) and the level properties $E_0\varphi_c = E_0\phi = \alpha$, we get

$$\begin{aligned} 0 &\leq \int_{\mathcal{X}} [\varphi_c(x) - \phi(x)][p_1(x) - cp_0(x)]\mu(dx) \\ &= E_1\varphi_c - E_1\phi - c[E_0\varphi_c - E_0\phi] \\ &= E_1\varphi_c - E_1\phi. \end{aligned}$$

Hence $E_1\varphi_c \geq E_1\phi$.

Finally, let φ be best at level $\alpha(\varphi) = \alpha \in (0, 1)$. But, $\psi_\alpha(L_{0,1})$ is another test with size α. By (2) of this lemma we know that it $\psi_\alpha(L_{0,1})$ is best too, i.e $E_1\psi_\alpha(L_{0,1}) = E_1\varphi$.
Hence, since $E_0\phi = E_0\psi_\alpha(L_{0,1})$,

$$\begin{aligned} 0 &= E_1\psi_\alpha(L_{0,1}) - E_1\varphi \\ &= E_1\psi_\alpha(L_{0,1}) - E_1\varphi - c_{1-\alpha}[E_0\psi_\alpha - E_0\varphi] \\ &= \int_{\mathcal{X}} [(\psi_\alpha \circ L_{0,1})(x) - \varphi(x)][p_1(x) - c_{1-\alpha}p_0(x)]\mu(dx), \end{aligned}$$

the second equality following from $E_0\varphi = \alpha = E_0\psi_\alpha$. Since

$$[(\psi_\alpha \circ L_{0,1})(x) - \varphi(x)][p_1(x) - c_{1-\alpha}p_0(x)] \geq 0 \qquad \text{for all } x \in \mathcal{X},$$

the previous two displays yield that

$$[(\psi_\alpha \circ L_{0,1})(x) - \varphi(x)][p_1(x) - c_{1-\alpha}p_0(x)] = 0 \quad \text{for } \mu\text{-almost all } x \in \mathcal{X}.$$

Thus, we have that

$$\mu\left(\left\{(\psi_\alpha \circ L_{0,1})(x) - \phi(x) \neq 0\right\} \cap \left\{p_1(x) - c_{1-\alpha}p_0(x) \neq 0\right\}\right) = 0,$$

that is, outside $\{p_1(x) - c_{1-\alpha}p_0(x) = 0\}$ it holds for μ-almost all $x \in \mathcal{X}$ that $(\psi_\alpha \circ L_{0,1})(x) - \phi(x) = 0$.

$\square$

Unbiasedness of tests

Definition 10

A level $\alpha \in (0,1)$ test φ is called *unbiased* for testing

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

in the experiment $\big(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\big)$ if $\inf_{\theta \in \Theta_1} E_\theta \varphi \geq \alpha$.

- In words, the rejection probability under the alternative is never smaller than that under the null.

Unbiasedness of LR test in binary experiment

Corollary 11 (Unbiasedness of LR test in binary experiment)

Let $\alpha \in (0,1)$ and consider testing $H_0 : P_0$ against $H_1 : P_1$ in $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_0, P_1\}\big)$, where P_0 and P_1 are dominated by a σ-finite measure μ. Then the LR test φ_c with $\alpha(\varphi_c) \in (0,1)$ satisfies $E_0\varphi_c \le E_1\varphi_c$, where the inequality is strict whenever $P_0 \neq P_1$.

- It follows that in particular $\alpha = E_0\psi_\alpha \le E_1\psi_\alpha$, with strict inequality whenever $P_0 \neq P_1$.
- Thus, the LR tests are *unbiased* since their rejection probability under $H_0 : P_0$ is not greater than under $H_1 : P_1$.
- Indeed the rejection probability under the null is strictly less than under a distinct alternative.

Proof

Abbreviate $\alpha = \alpha(\varphi_c)$. Clearly, the test $\phi_\alpha \equiv \alpha$ is of size α and satisfies $E_1\phi_\alpha = \alpha$. By the Neyman-Pearson Lemma (part 2), φ_c is most powerful against P_1. Hence, $E_1\varphi_c \geq \alpha = E_0\varphi_c$.

Let $P_0 \neq P_1$ and assume for contradiction that $E_1\varphi_c = \alpha$. Now,

$$
\begin{aligned}
&\left\{p_1(x) - c_{1-\alpha}p_0(x) \neq 0\right\} \\
\subseteq &\left\{\psi_\alpha(x) - \phi_\alpha(x) \neq 0\right\} \cap \left\{p_1(x) - c_{1-\alpha}p_0(x) \neq 0\right\},
\end{aligned}
$$

since $\psi_\alpha(L_{0,1})$ is an LR test and thus equals 0 or 1 when $p_1 - c_{1-\alpha}p_0 \neq 0$; but $\phi_\alpha \equiv \alpha \in (0,1)$ and hence $\psi_\alpha - \phi_\alpha \neq 0$. Since $E_1\varphi_c = \alpha$ (assumed for contradiction), then actually ϕ_α is most powerful. By the Neyman-Pearson Lemma (part 3)

$$\mu(p_1(x) - c_{1-\alpha}p_0(x) \neq 0) = 0.$$

Thus, $p_1 = c_{1-\alpha} p_0$ for μ-almost all $x \in \mathcal{X}$. If $c_{1-\alpha} \neq 1$ then p_1 and p_0 can't integrate to 1; a contradiction. If, on the other hand, $c_{1-\alpha} = 1$, then $p_0 = p_1$ μ-a.e. and hence $P_0 = P_1$, which can't be since $P_0 \neq P_1$.

$\square$

Examples: Underlying common strucure

- Let us study some examples of likelihood ratio tests for concrete testing problems.
- In finding an LR test of size $\alpha \in (0, 1)$ we follow the recipe indicated around (3).
  1. Find $F_0^{-1}(1 - \alpha)$.
  2. Write down γ_α, the value of the test on the randomization set,

$$\gamma_\alpha = \begin{cases} \frac{F_0(c_{1-\alpha})-(1-\alpha)}{P_0(L_{0,1}=c_{1-\alpha})} & \text{if } P_0(L_{0,1} = c_{1-\alpha}) > 0 \\ 0 & \text{if } P_0(L_{0,1} = c_{1-\alpha}) = 0. \end{cases}$$

     No randomization is needed when $L_{0,1}$ has a continuous cdf under P_0 as then $P_0(L_{0,1} = c_{1-\alpha}) = 0$.
  3. Write down the most powerful test at level α:

$$\psi_\alpha(L_{0,1}) = \gamma_\alpha \mathbb{1}_{\{L_{0,1}=c_{1-\alpha}\}} + \mathbb{1}_{\{L_{0,1}>c_{1-\alpha}\}}$$

Example: Normal distribution

- Consider testing $P_0 = \mathrm{N}(\mu_0, \sigma^2)$ against $P_1 = \mathrm{N}(\mu_1, \sigma^2)$.
- Recall (pages 28, 33) that the LR of P_1 with respect to P_0 is

$$L_{0,1}(x) = \frac{p_1(x)}{p_0(x)} = e^{\frac{x(\mu_1 - \mu_0)}{\sigma^2} - \frac{\mu_1^2 - \mu_0^2}{2\sigma^2}}.$$

- For all $c \in [0, \infty)$ one has $P_0\left(L_{0,1} = c\right) = 0$, such that no randomization is needed to achieve desired size.
- For $\alpha \in (0,1)$ there exists a $c_{1-\alpha}$ such that

$$F_0(c_{1-\alpha}) = 1 - \alpha.$$

- Since $P_0(L_{0,1} = c_{1-\alpha}) = 0$ we have $\gamma_\alpha = 0$.
- By the Neyman-Pearson Lemma, the test $\mathbb{1}_{\{L_{0,1} > c_{1-\alpha}\}}$ is thus most powerful against P_1 at level α, cf. the construction in (3) .

Example: Uniform distribution

We shall see that here the randomization set has P_0-probability one!

Consider testing $H_0 : P_0 = \mathrm{U}(0,1)$ against $H_1 : P_1 = \mathrm{U}(0,\theta)$ for $\theta > 1$ and recall (page 34) that

$$L_{0,1}(x) = \frac{1}{\theta}\mathbb{1}_{[0,1]}(x) + \infty \cdot \mathbb{1}_{(1,\theta]}(x).$$

Since $P_0(L_{0,1} = \infty) = 0$ always, one observes that for any $c \in [0,\infty)$

$$\begin{aligned} F_0(c) &= P_0\left(L_{0,1} \leq c\right) \\ &= P_0\left(\{L_{0,1} \leq c\} \cap \{L_{0,1} < \infty\}\right) \\ &= P_0\left(x \in \mathcal{X} : \frac{1}{\theta}\mathbb{1}_{[0,1]}(x) \leq c\right) \\ &= \begin{cases} 0 & \text{if } c < \frac{1}{\theta} \\ 1 & \text{if } c \geq \frac{1}{\theta}. \end{cases} \end{aligned}$$

Hence, for all $\alpha \in (0,1)$ one observes that

$$c_{1-\alpha} = F_0^{-1}(1-\alpha) = \frac{1}{\theta},$$

and

$$P_0(L_{0,1} = c_{1-\alpha}) = 1,$$

that is the randomization set has P_0-probability 1!

Clearly, $F_0(c_{1-\alpha}) = P_0(L_{0,1} \leq c_{1-\alpha}) = 1$. Thus, in accordance with the definition of ψ_α in (3) we set

$$\gamma_\alpha = \frac{F_0(c_{1-\alpha}) - (1-\alpha)}{P_0(L_{0,1} = c_{1-\alpha})} = \alpha.$$

It follows from the Neyman-Pearson Lemma that

$$\psi_\alpha(L_{0,1}) = \alpha \cdot \mathbb{1}_{\{L_{0,1}=1/\theta\}} + \mathbb{1}_{\{L_{0,1}>1/\theta\}}$$

is most powerful against P_1 at level α. Equivalently,

$$\psi_\alpha(L_{0,1}(x)) = \alpha \cdot \mathbb{1}_{[0,1]}(x) + \mathbb{1}_{(1,\theta]}(x) \tag{5}$$

is most powerful against P_1 at level α.

Monotone Likelihood Ratios

- So far we have mainly studied testing a simple null P_0 against a simple alternative P_1.
- This is somewhat restrictive since problems arising in applications typically involve (parametric) *families* of distributions and not just two single ones.
- The Neyman-Pearson Lemma can be used to establish the existence of uniformly most powerful tests in certain one-sided testing problems.

We now again consider an experiment

$$\mathcal{E} = \left(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\right)$$

with $\Theta \subseteq \mathbb{R}$. That is "a one-dimensional parameter".

Monotone Likelihood Ratio: MLR

- In the following we write $L_{\theta_0,\theta_1} := \frac{dP_{\theta_1}}{dP_{\theta_0}}$.

Definition 12 (Monotone Likelihood Ratio: MLR)

An experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\big)$ with $\Theta \subseteq \mathbb{R}$ is said to have a nondecreasing (increasing) likelihood ratio in the statistic $T : \mathcal{X} \mapsto \mathbb{R}$ if for every $\theta_0, \theta_1 \in \Theta$ with $\theta_0 < \theta_1$ there is a nondecreasing (increasing) function $h_{\theta_0,\theta_1} : \mathbb{R} \to [0, \infty]$ such that $L_{\theta_0,\theta_1} = h_{\theta_0,\theta_1}(T)$, $\{P_{\theta_0}, P_{\theta_1}\}$-almost surely. We also say in short that $\{P_\theta : \theta \in \Theta\}$ has (strict) MLR, i.e. (strictly) monotone likelihood ratio in T.

Example: One-parameter exponential family

- Consider a one-parameter exponential family in canonical form; $\{P_\eta : \eta \in \Delta\}$, $\Delta \subseteq \mathbb{R}$ being the natural parameter space.
- By convexity of Δ, it is an interval.
- Let μ be a dominating measure of the family and

$$\frac{dP_\eta}{d\mu} = e^{\eta T(x) - B(\eta)} h(x).$$

- Then, letting $\eta_0 < \eta_1$ and using that P_{η_0} and P_{η_1} are mutually absolutely continuous one has by (2) on page 27

$$\frac{dP_{\eta_1}}{dP_{\eta_0}} = e^{(\eta_1 - \eta_0)T(x) - B(\eta_1) + B(\eta_0)} \cdot \mathbb{1}_{\left\{\frac{dP_{\eta_0}}{d\mu} > 0\right\}},$$

- Since $\frac{dP_{\eta_0}}{d\mu} = 0$ only if $h = 0$ and $P_{\eta_0}(h = 0) = 0$, we can in fact use

$$\frac{dP_{\eta_1}}{dP_{\eta_0}} = e^{(\eta_1 - \eta_0)T(x) - B(\eta_1) + B(\eta_0)}$$

  as a density $\frac{dP_{\eta_1}}{dP_{\eta_0}}$ is only unique up to P_{η_0}-null sets.
- Now, for $\eta_1 > \eta_0$,

$$\frac{dP_{\eta_1}}{dP_{\eta_0}} = e^{(\eta_1 - \eta_0)T(x) - B(\eta_1) + B(\eta_0)}$$

  is clearly increasing in (the sufficient statistic) T.
- Thus, the family $\{P_\eta : \eta \in \Delta\}$ has strict MLR.
  [use $h_{\eta_0,\eta_1}(t) = e^{(\eta_1 - \eta_0)t - B(\eta_1) + B(\eta_0)}$ which is increasing in t.]

Example: continued

- If, furthermore, Θ is an interval and $\eta : \Theta \to \Delta$ is nondecreasing (increasing), then $\{P_{\eta(\theta)} : \theta \in \Theta\}$ has (strict) MLR in T.
- Here it is of course useful to think of η as mapping the "structural/statistically interpretable" parameters to the natural parameter space.

$$\frac{dP_{\theta_1}}{dP_{\theta_0}} = e^{(\eta(\theta_1)-\eta(\theta_0))T(x)-B(\eta(\theta_1))+B(\eta(\theta_0))}$$

- If $\theta_1 > \theta_0 \Rightarrow \eta(\theta_1) \geq \eta(\theta_0)$ then $\{P_{\eta(\theta)} : \theta \in \Theta\}$ has a monotone likelihood ratio in T.
- We have seen examples of these mappings when we introduced exponential families.
- Specifically, we studied the normal and Poisson families of distributions.

Details for Poisson distribution

The Poisson distribution has density

$$p_\lambda(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad \lambda \in (0,\infty)$$

with respect to the counting measure τ on $(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$, cf. the exercises. Clearly,

$$p_\lambda(x) = \exp\left(\log(\lambda)x - \lambda\right)\frac{1}{x!},$$

such that $m = 1$, $\eta_1(\lambda) = \log(\lambda)$, $T_1(x) = x$, $B(\lambda) = \lambda$ and $h(x) = \frac{1}{x!}$.

Since $\lambda \mapsto \eta_1(\lambda)$ is strictly increasing, we conclude that

$$\left(\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0), \{\mathrm{Poi}(\lambda), \lambda \in (0,\infty)\}\right)$$

has a strictly increasing likelihood ratio in $T(x) = x$.

MLR outside exponential families: Uniform distribution

- Consider observing n independent observations $X_1, \ldots, X_n$ from a member of $\{U[0, \theta] : \theta > 0\}$.
- Then, for $\theta_0 < \theta_1$, by the usual arguments, $M := \max(X_1, \ldots, X_n)$,

$$L_{\theta_0,\theta_1} = (\theta_0/\theta_1)^n \mathbb{1}_{[0,\theta_0]}(M) + \infty \cdot \mathbb{1}_{(\theta_0,\theta_1]}(M),$$

  which is nondecreasing in M, $\{P_{\theta_0}, P_{\theta_1}\}$-almost surely. [Note that L_{θ_0,θ_1} drops down to 0 after θ_1, but this is a $\{P_{n,\theta_0}, P_{n,\theta_1}\}$-null set where $P_{n,\theta} = \otimes_{i=1}^n U[0, \theta]$ is the measure for the n-sample.]

A remark on uniform distributions

- Observe that the uniform family of distributions is a typical example of when we can "get things" outside exponential families.
- We used it as an example for existence of
  1. sufficient statistics
  2. complete statistics
  3. MLR

  outside exponential families.

In the experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\big)$, with $\Theta \subseteq \mathbb{R}$, we wish to test

$$H_0 : \theta \in \Theta_0 = (-\infty, \theta_0] \cap \Theta \quad \text{vs.} \quad H_1 : \theta \in \Theta_1 = (\theta_0, \infty) \cap \Theta$$

for some $\theta_0 \in \Theta$, where $\Theta_0, \Theta_1 \neq \emptyset$.

In this *one-sided* testing problem we shall now construct a UMP test.

Karlin-Rubin Theorem: Preliminary definitions

Assume that $P_{\theta_0}(T < \infty) = 1$ and let $F_{\theta_0}(t) = P_{\theta_0}(T \leq t)$, where T is the statistic in which the likelihood ratio is monotone.

With $c_{1-\alpha} := F_{\theta_0}^{-1}(1-\alpha)$ for $\alpha \in (0,1)$, define

$$\psi_\alpha(t) := \begin{cases} 1 & \text{if } t > c_{1-\alpha} \\ \gamma_\alpha & \text{if } t = c_{1-\alpha}, \\ 0 & \text{if } t < c_{1-\alpha} \end{cases} \qquad t \in [0, \infty]$$

where

$$\gamma_\alpha = \begin{cases} \frac{F_{\theta_0}(c_{1-\alpha}) - (1-\alpha)}{P_{\theta_0}(T = c_{1-\alpha})} & \text{if } P_{\theta_0}(T = c_{1-\alpha}) > 0 \\ 0 & \text{if } P_{\theta_0}(T = c_{1-\alpha}) = 0. \end{cases}$$

Define

$$\varphi_{T,\alpha}(x) := \psi_\alpha(T(x)), \quad x \in \mathcal{X}. \tag{6}$$

- Importantly, $\varphi_{T,\alpha}$ does not depend on other parameters than θ_0 (which is fixed).

Minor remarks:

- The construction of $\varphi_{T,\alpha} = \psi_\alpha(T)$ is of course identical to that of $\psi_\alpha(L_{0,1})$ in the context of LR tests in binary experiments.
- The cdf F_0 of $L_{0,1}$ under P_0 and ensuing quantities are replaced by the cdf F_{θ_0} of T under P_{θ_0} and their ensuing quantities.

Karlin-Rubin Theorem

Theorem 13 (Karlin-Rubin)

Suppose that in the experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\big)$, *with* $\Theta \subseteq \mathbb{R}$, *the family of distributions has a nondecreasing likelihood ratio in the statistic* $T : \mathcal{X} \to \mathbb{R}$. *Let* $H_0 : \theta \in \Theta_0 = (-\infty, \theta_0] \cap \Theta$ *and* $H_1 : \theta \in \Theta_1 = (\theta_0, \infty) \cap \Theta$ *for some* $\theta_0 \in \Theta$, *where* $\Theta_0, \Theta_1 \neq \emptyset$. *For* $\alpha \in (0,1)$ *the test* $\varphi_{T,\alpha}$ *in* (6) *is uniformly most powerful at level* α *for* H_0 *versus* H_1. *Furthermore,*

1. $\theta \mapsto E_\theta \varphi_{T,\alpha}$ *is nondecreasing.*
2. $E_{\underline{\theta}} \varphi_{T,\alpha} \leq \alpha = E_{\theta_0} \varphi_{T,\alpha} \leq E_{\overline{\theta}} \varphi_{T,\alpha}, \quad \underline{\theta} \leq \theta_0 \leq \overline{\theta}, \quad \underline{\theta}, \overline{\theta} \in \Theta.$
3. *For every test* φ *of level* α *it holds that*

$$E_{\overline{\theta}} \varphi_{T,\alpha} \geq E_{\overline{\theta}} \varphi, \qquad \overline{\theta} > \theta_0.$$

Proof

By the same argument as in the proof of part 1 of the Neyman-Pearson Lemma, one has that $E_{\theta_0}\varphi_{T,\alpha} = \alpha$.

To see that $\varphi_{T,\alpha}$ is UMP, fix $\theta_1 > \theta_0$. By the monotone likelihood ratio property there exists a $h_{\theta_0,\theta_1} : \mathbb{R} \to [0, \infty]$ which is nondecreasing and such that $L_{\theta_0,\theta_1} = h_{\theta_0,\theta_1}(T)$, $\{P_{\theta_0}, P_{\theta_1}\}$-almost surely. With $c_{1-\alpha} := F_{\theta_0}^{-1}(1-\alpha)$ define $c := h_{\theta_0,\theta_1}(c_{1-\alpha})$.

Observe that

$$\begin{aligned} T &> c_{1-\alpha} \quad \text{if } L_{\theta_0,\theta_1} > c \\ T &< c_{1-\alpha} \quad \text{if } L_{\theta_0,\theta_1} < c \end{aligned}$$

and thus

$$\begin{aligned} \varphi_{T,\alpha} &= 1 \quad \text{if } L_{\theta_0,\theta_1} > c \\ \varphi_{T,\alpha} &= 0 \quad \text{if } L_{\theta_0,\theta_1} < c. \end{aligned}$$

It follows that $\varphi_{T,\alpha}$ is a likelihood ratio test with $E_{\theta_0}\varphi_{T,\alpha} = \alpha$ and thus is most powerful for testing the simple hypothesis $H_0 : P_{\theta_0}$ against $H_1 : P_{\theta_1}$ by the Neyman-Pearson Lemma.

Now $\varphi_{T,\alpha}$ does not depend on the specific value of $\theta_1 > \theta_0$, cf. the definition of $\varphi_{T,\alpha}$. Thus, the above argument can be repeated for any $\theta > \theta_0$.

It follows that $\varphi_{T,\alpha}$ is most powerful in testing P_{θ_0} against any P_θ with $\theta > \theta_0$.

We next show that $\theta \mapsto E_\theta\varphi_{T,\alpha}$ is nondecreasing from which it will follow that

$$\sup_{\theta \in \Theta_0} E_\theta\varphi_{T,\alpha} = E_{\theta_0}\varphi_{T,\alpha} = \alpha,$$

that is $\varphi_{T,\alpha}$ has size α.

To see that $\theta \mapsto E_\theta \varphi_{T,\alpha}$ is nondecreasing, let $\underline{\theta} < \overline{\theta}$. $\varphi_{T,\alpha}$ is an LR test for $\underline{\theta}$ versus $\overline{\theta}$ by the same arguments as above (simply replace θ_0 by $\underline{\theta}$ and θ_1 by $\overline{\theta}$) with size $E_{\underline{\theta}}\varphi_{T,\alpha}$. It follows from unbiasedness of LR tests, cf. Corollary 11, that

$$E_{\underline{\theta}}\varphi_{T,\alpha} \leq E_{\overline{\theta}}\varphi_{T,\alpha},$$

which establishes (1) and (2) of the theorem. In particular $\varphi_{T,\alpha}$ has size α as

$$E_{\underline{\theta}}\varphi_{T,\alpha} \leq E_{\theta_0}\varphi_{T,\alpha} = \alpha \qquad \text{for all} \qquad \underline{\theta} \leq \theta_0.$$

cf. the first inequality in (2).

Now let φ be any other test of level α, that is

$$\sup_{\theta \in \Theta_0} E_\theta \varphi \leq \alpha.$$

Then, in particular, $E_{\theta_0}\varphi \leq \alpha$. Now since $\varphi_{T,\alpha}$ is most powerful in any of the binary testing problems $H_0 : P_{\theta_0}$ against $H_1 : P_\theta$ with $\theta > \theta_0$, it follows that

$$E_\theta \varphi_{T,\alpha} \geq E_\theta \varphi \qquad \text{for all } \theta > \theta_0.$$

Because φ was an arbitrary test of level α, it follows that $\varphi_{T,\alpha}$ is uniformly most powerful at level α.

$\square$

Normal distribution

- Consider the family $\{N(\mu, \sigma_0^2) : \mu \in \mathbb{R}\}$, where $\sigma_0^2 > 0$ is known (recall Karlin-Rubin is for one-dimensional parameters).
- Based on n independent observations $X_1, \ldots, X_n$ we wish to test

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0,$$

for some $\mu_0 \in \mathbb{R}$.
- We have seen that this is an exponential family with $\eta(\mu) = \mu/\sigma_0^2$.
- Since $\mu \mapsto \mu/\sigma_0^2$ is strictly increasing, we have the strict MLR property in the sufficient statistic $T = \sum_{i=1}^n X_i$.
- Thus, by the Karlin-Rubin theorem, the test in (6)

$$\psi_\alpha(T) := \mathbb{1}_{(c_{1-\alpha},\infty)}(T),$$

where we use that $P_{\theta_0}(T = c_{1-\alpha}) = 0$ such that $\gamma_\alpha = 0$, is UMP for H_0 versus H_1.

- Let us determine $c_{1-\alpha}$ (which is chosen according to the construction around (6)).
- For brevity we introduce $P_\mu = \mathrm{N}(\mu, \sigma_0^2)$.
- Note that $P_{\mu_0}^n \circ T^{-1} = \mathrm{N}(n\mu_0, n\sigma_0^2)$.
- Hence, defining $F_{n,\mu_0}(t) := (P_{\mu_0}^n \circ T^{-1})(-\infty, t]),\ t \in \mathbb{R}$ one has

$$c_{1-\alpha} := F_{n,\mu_0}^{-1}(1-\alpha) = z_{1-\alpha}\sqrt{n}\sigma_0 + n\mu_0, \tag{7}$$

where $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$ and $\alpha \in (0,1)$.
- Hence, defining $Z_0 := \frac{T - n\mu_0}{\sqrt{n}\sigma_0}$, the UMP test is

$$\mathbb{1}_{(z_{1-\alpha}\sqrt{n}\sigma_0 + n\mu_0, \infty)}(T) = \mathbb{1}_{(z_{1-\alpha}, \infty)}(Z_0),$$

which is nothing else than the classic "Z-test"!
- Hence, the Z-test is UMP for $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$.

Power of Z-test

- The Z-test is UMP, but what is its power against a specific $\mu_1 \neq \mu_0$?
- Fix, $\mu_1 \in \mathbb{R}$ and note that

$$
\begin{aligned}
E_{n,\mu_1} \mathbb{1}_{(z_{1-\alpha},\infty)}(Z_0) &= P^n_{\mu_1}(Z_0 > z_{1-\alpha}) \\
&= P^n_{\mu_1}\left(\frac{T - n\mu_1}{\sqrt{n}\sigma_0} > z_{1-\alpha} + \frac{\sqrt{n}(\mu_0 - \mu_1)}{\sigma_0}\right) \\
&= 1 - \Phi\left(z_{1-\alpha} + \frac{\sqrt{n}(\mu_0 - \mu_1)}{\sigma_0}\right). \qquad (8)
\end{aligned}
$$

1. If $\mu_1 > \mu_0$, then power tends to 1 as $n \to \infty$.
2. If $\mu_1 = \mu_0$, then power is constant at $1 - \Phi(z_{1-\alpha}) = \alpha$.
3. If $\mu_1 < \mu_0$, then power tends to 0 as $n \to \infty$.

- Observe also that for every $n \in \mathbb{N}$ the power function $\mu \mapsto E_{n,\mu} \mathbb{1}_{(z_{1-\alpha},\infty)}(Z_0)$ is increasing.
- This is no coincidence as the likelihood ratio is increasing.

Further remarks

- We repeat that for any fixed $\mu_1 > \mu_0$, the the power in (8) is the highest power that *any* level α test can achieve.
- This is a strong justification for using the "Z-test".
- In accordance with the Karlin-Rubin Theorem, the power function is increasing.

Theorem 14 (UMP tests in exponential families)

Suppose that $\{P_{\eta(\theta)} : \theta \in (a, b)\}$ *is a one-parameter exponential family with* μ*-density*

$$\frac{dP_\theta}{d\mu}(x) = e^{\eta(\theta)T(x) - B(\eta(\theta))}h(x),$$

and $\theta \mapsto \eta(\theta)$ *non-decreasing. Then for* $\theta_0 \in (a, b)$ *and testing*

$$H_0 : \theta \in (a, \theta_0] \quad vs. \quad H_1 : \theta \in (\theta_0, b)$$

the test [recall details of notation above equation (6)]

$$\varphi_{T,\alpha} = \gamma_\alpha \mathbb{1}_{\{c_{1-\alpha}\}}(T) + \mathbb{1}_{(c_{1-\alpha},\infty)}(T)$$

is uniformly most powerful, where $c_{1-\alpha} = F_{\theta_0}^{-1}(1 - \alpha)$.

This result is "obvious" from the Karlin-Rubin Theorem since under the stated assumptions $\{P_{\eta(\theta)} : \theta \in (a, b)\}$ has a monotone likelihood ratio.

UMP is based on complete sufficient statistic in exp fam

- Recall that in exponential families the statistic T that the LR is monotone in, is also generally the complete sufficient statistic.
- Thus, in the uniformly most powerful test is based on this.
- That is, the test $\varphi_{T,\alpha}$ only depends on the data via $T(x)$.

A partial converse

- Pfanzagl (1968) showed that under mild conditions, the existence of a uniformly best level α test for one-sided alternatives, for one $\alpha \in (0,1)$ and all sample sizes $n \in \mathbb{N}$ implies that the underlying family of distributions is an exponential family.
- Thus, apart from a few exceptions (e.g. uniform distributions, cf. next slide) the theory of one-sided UMP tests is one of exponential families.

Example: Uniform distribution

- Consider observing n independent observations $X_1, \ldots, X_n$ from a member of $\{\mathrm{U}[0,\theta] : \theta > 0\}$.
- For brevity we shall sometimes write $P_\theta = \mathrm{U}[0,\theta]$.
- We have seen that this family of distributions has monotone likelihood ratio in $M := \max(X_1, \ldots, X_n)$.
- Hence, in testing (for fixed $\theta_0 > 0$)

  $$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0.$$

  the test

  $$\psi_\alpha(M) := \gamma_\alpha \mathbb{1}_{\{c_{1-\alpha}\}}(M) + \mathbb{1}_{(c_{1-\alpha},\infty)}(M)$$

  is UMP where $c_{1-\alpha} = F_{n,\theta_0}^{-1}(1-\alpha)$ with F_{n,θ_0} being the cdf of $P_{\theta_0}^n \circ M^{-1}$.

- We have seen in the estimation slides that

$$F_{n,\theta_0}(t) = \left(\frac{t}{\theta_0}\right)^n \qquad \text{for } t \in (0, \theta_0),$$

while $F_{n,\theta_0}(t) = 0$ for $t \leq 0$ and $F_{n,\theta_0}(t) = 1$ for $t \geq \theta_0$.
- Hence, for $\alpha \in (0, 1)$,

$$c_{1-\alpha} = \theta_0 \cdot (1-\alpha)^{1/n} \in (0, \theta_0).$$

- Observe that $P^n_{\theta_0}(M = c_{1-\alpha}) = 0$ (since M has a continuous distribution).
- Hence, the UMP test is

$$\psi_\alpha(M) := \mathbb{1}_{(c_{1-\alpha},\infty)}(M).$$

- We start rejecting at the upper end of the support of the null distribution P_{θ_0} as we reject for $M > \theta_0 \cdot (1-\alpha)^{1/n} \in (0, \theta_0)$.

Comparison to the LR test

Recall that in case $\theta_0 = 1$ and $n = 1$, such that $M = X$, we have seen in equation (5) on page 53, that the most powerful test of level $\alpha \in (0, 1)$ for testing the simple null and alternative

$$H_0 : \theta = 1 \qquad \text{vs.} \qquad H_1 : \theta = \theta_1,$$

for some $\theta_1 > 1$ is

$$\psi_{LR,\alpha}(X) := \alpha \mathbb{1}_{[0,1]}(X) + \mathbb{1}_{(1,\infty)}(X).$$

Of course the UMP test of level α for testing

$$H_0' : \theta \leq 1 \qquad \text{vs.} \qquad H_1' : \theta > 1$$

that we have just constructed on the previous slide is (as $c_{1-\alpha} = 1 - \alpha$ for $n = 1$ and $\theta_0 = 1$)

$$\psi_\alpha(X) := \mathbb{1}_{(c_{1-\alpha},\infty)}(X) = \mathbb{1}_{(1-\alpha,\infty)}(X).$$

- Since ψ_α is *uniformly* most powerful for H_0' vs. H_1' it must in particular be most powerful against θ_1.
- Also, in the simple testing problem H_0 vs. H_1, the LR test $\psi_{LR,\alpha}$ is most powerful.
- But $\psi_{LR,\alpha}$ and ψ_α do not "look identical".
    - Recall that by the Neyman-Pearson Lemma, the most powerful test of level α is only uniquely determined outside the randomization set $\{L_{1,\theta_1} = c_{1-\alpha}\} = \{X \in [0,1]\}$.
- Is there a contradiction here?
- Of course not!
- Outside the set $\{X \in [0,1]\}$ the two tests agree as both reject for $X > 1$.

Of course $\psi_{LR,\alpha}$ and ψ_α must also have the same power against θ_1 as otherwise they could not both be most powerful.

But this is the case since

$$E_{\theta_1}\psi_{LR,\alpha} = \alpha P_{\theta_1}([0,1]) + P_{\theta_1}((1,\infty)) \\ = \frac{\alpha}{\theta_1} + 1 - \frac{1}{\theta_1} = 1 - \frac{1-\alpha}{\theta_1}$$

and

$$E_{\theta_1}\psi_\alpha = P_{\theta_1}((1-\alpha,\infty)) = 1 - \frac{1-\alpha}{\theta_1}.$$

Note that it is in accordance with intuition that the power is increasing in θ_1 as then P_1 and P_{θ_1} are further from each other.

Finding $c_{1-\alpha}$

In order to find $c_{1-\alpha}$ we used that the distribution of T under P_{θ_0} was known. The distribution of T under P_{θ_0} may not always be that easily available, in which case one may have to resort to numerical solutions to find its $(1-\alpha)$-percentile $c_{1-\alpha}$.

$P_{\theta_0}(T = c_{1-\alpha}) = 0$: The UMP test becomes non-randomized

If $P_{\theta_0}(T = c_{1-\alpha}) = 0$ the best uniformly best level α test becomes nonrandomized ($\gamma_\alpha = 0$). As the above examples illustrate, this occurs when the cdf of T is continuous under P_{θ_0}.

Karlin-Rubin Theorem: $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$

- So far we have studied UMP tests for $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$.
- However, one may also want to test $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$.
- Unsurprisingly, the situation is very similar.
- We only need to modify $\varphi_{T,\alpha}$ slightly.
- Observe that we now reject for "small values of T" in the construction on the next slide.

Adjusting $\varphi_{T,\alpha}$ to $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$

Let $F_{\theta_0}(t) = P_{\theta_0}(T \leq t)$.

With $c_\alpha := F_{\theta_0}^{-1}(\alpha)$ for $\alpha \in (0,1)$, define

$$\psi_\alpha(t) := \begin{cases} 1 & \text{if } t < c_\alpha \\ \gamma_\alpha & \text{if } t = c_\alpha, \\ 0 & \text{if } t > c_\alpha \end{cases} \qquad t \in [0, \infty]$$

where

$$\gamma_\alpha = \begin{cases} \frac{\alpha - P_{\theta_0}(T < c_\alpha)}{P_{\theta_0}(T = c_\alpha)} & \text{if } P_{\theta_0}(T = c_\alpha) > 0 \\ 0 & \text{if } P_{\theta_0}(T = c_\alpha) = 0. \end{cases}$$

Define

$$\varphi_{T,\alpha}(x) := \psi_\alpha(T(x)), \quad x \in \mathcal{X}. \tag{9}$$

Karlin-Rubin Theorem: $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$

Let introduce a UMP for $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$.

Theorem 15 (Karlin-Rubin)

Suppose that in the experiment $\mathcal{E} = \big(\mathcal{X}, \mathcal{A}, \{P_\theta : \theta \in \Theta\}\big)$, *with* $\Theta \subseteq \mathbb{R}$, *the family of distributions has a nondecreasing likelihood ratio in the statistic* $T : \mathcal{X} \to \mathbb{R}$. *Let* $H_0 : \theta \in \Theta_0 = [\theta_0, \infty) \cap \Theta$ *and* $H_1 : \theta \in \Theta_1 = (-\infty, \theta_0) \cap \Theta$ *for some* $\theta_0 \in \Theta$, *where* $\Theta_0, \Theta_1 \neq \emptyset$. *For* $\alpha \in (0,1)$ *the test* $\varphi_{T,\alpha}$ *in* (9) *is uniformly most powerful at level* α *for* H_0 *versus* H_1. *Furthermore,*

1. $\theta \mapsto E_\theta \varphi_{T,\alpha}$ *is nonincreasing.*
2. $E_{\underline{\theta}} \varphi_{T,\alpha} \geq \alpha = E_{\theta_0} \varphi_{T,\alpha} \geq E_{\overline{\theta}} \varphi_{T,\alpha}, \quad \underline{\theta} \leq \theta_0 \leq \overline{\theta}, \quad \underline{\theta}, \overline{\theta} \in \Theta.$
3. *For every test* φ *of level* α *it holds that*

$$E_{\underline{\theta}} \varphi_{T,\alpha} \geq E_{\overline{\theta}} \varphi, \qquad \underline{\theta} < \theta_0.$$

"Generally" UMP tests do not exist

- The Karlin-Rubin theorem shows that in one-sided testing problems with MLR a UMP exists.
- However, the existence of a UMP is the exception rather than the norm.
- Typically, the identity of the most powerful test depends on the alternative.
- To illustrate this, consider testing, for $\mu_0 \in \mathbb{R}$,

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu < \mu_0 \text{ or } \mu > \mu_0$$

  within the one-parameter family $\{\mathrm{N}(\mu, \sigma_0^2) : \mu \in \mathbb{R}\}$ and $\sigma_0^2 > 0$ known.

- Let $\alpha \in (0, 0.5)$.
- By the Karlin-Rubin Theorem $\mathbb{1}_{(c_{1-\alpha},\infty)}(T)$ is most powerful against $\overline{\mu} > \mu_0$, where $T = \sum_{i=1}^n X_i$
- On the other hand, $\mathbb{1}_{(-\infty,c_\alpha)}(T)$ is most powerful against $\underline{\mu} < \mu_0$.
- Reasoning as around (7) on page 71 one gets that

$$c_\alpha = z_\alpha \sqrt{n}\sigma_0 + n\mu_0 < z_{1-\alpha}\sqrt{n}\sigma_0 + n\mu_0 = c_{1-\alpha}.$$

- Hence,

$$P_\mu^n \left(\mathbb{1}_{(c_{1-\alpha},\infty)}(T) \neq \mathbb{1}_{(-\infty,c_\alpha)}(T) \right) \geq P_\mu^n \left(T \geq c_{1-\alpha} \right) > 0,$$

for all $\mu \in \mathbb{R}$ and so there is no single test that is most powerful against $\underline{\mu}$ and $\overline{\mu}$. [can also just show this for a single $\mu \in \mathbb{R}$ and then appeal to mutual absolute continuity.]

- One reaction to the typical non-existence of UMP tests is to further restrict the class of tests.
- For example, it can be shown that within the class of *unbiased* tests (Definition 10, page 45) a UMP test *does* exist for

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu < \mu_0 \text{ or } \mu > \mu_0.$$

- Observe that unbiasedness rules out the one-sided tests $\mathbb{1}_{(-\infty, c_\alpha)}(T)$ and $\mathbb{1}_{(c_{1-\alpha}, \infty)}(T)$ above, cf. also the power function of $\mathbb{1}_{(c_{1-\alpha}, \infty)}$ in (8).
- Unbiasedness seems reasonable to require of a test.
- If, on the other hand a test is only slightly biased against uninteresting alternatives (close to the null) but much more powerful against interesting alternatives then imposing unbiasedness is more questionable, cf. Pfanzagl (1994) page 127.
- See Chapter 4 in Lehmann and Romano (2005) for more properties on unbiased tests.

Summary

- We introduced the general testing problem and tests.
- We argued that in general no test exists that simultaneously minimizes the probability of Type 1 and Type 2 error.
- We established the Neyman-Pearson Lemma, which gives the structure of the most powerful level $\alpha \in (0, 1)$ tests in binary testing problems.
    - This most powerful test is a likelihood ratio test.
- We established the Karlin-Rubin Theorem on UMP tests in one-sided one-parameter testing problems.

References

Lehmann, E. and J. Romano (2005): *Testing Statistical Hypotheses*, Springer.

Liese, F. and K.-J. Miescke (2008): *Statistical Decision Theory*, Springer.

Pfanzagl, J. (1968): "A characterization of the one parameter exponential family by existence of uniformly most powerful tests," *Sankhyā: The Indian Journal of Statistics, Series A*, 147–156.

——— (1994): *Parametric Statistical Theory*, Walter de Gruyter.

Witting, H. (1985): *Mathematische Statistik I: Parametrische Verfahren bei festem Stichprobenumfang*, Springer.