Hypothesis Testing

STAT3023

Rachel Wang

School of Mathematics and Statistics, USyd

14 Sep, 2022

General setup and definitions

- Examples of statistical hypothesis testing problems:
  - decide based on clinical trial data (control vs. treatment) whether a new drug lowers blood pressure
  - decide if the life-time of a mechanical component in a car follows an exponential distribution or another distribution
- Treat observed data as values taken by random variables (often assumed to be iid)
- Statistical hypothesis: an assertion or conjecture about the underlying **distribution** of the random variables
- Null hypothesis H_0 and alternative hypothesis H_1: a distribution (or a family of distributions) to be compared

General setup and definitions

- Test statistic: a real-valued function $T(\mathbf{X})$ of the data $\mathbf{X} = (X_1, \ldots, X_n)$, capturing certain features of the distribution.
- P-value: assuming H_0 true,

$$P(\text{at least as much evidence against } H_0 \text{ as was observed}),$$

(or $\sup_{P \in H_0} P(...)$). The smaller the p-value, the stronger the evidence against H_0.

General setup and definitions

- Goal: **compare different tests**
- A test procedure partitions possible values of $\mathbf{X} = (X_1, \ldots, X_n)$ into two subsets: an acceptance region and a **rejection region (or critical region)**, denoted C, assuming H_0:

Reject H_0 if $\underset{\sim}{X} \in C$

Accept H_0 if $\underset{\sim}{X} \in C^c$

Two types of errors

- Type I and Type II errors

	H_0 is true	H_0 is false
Accept H_0	no error	Type II $= P_{H_1}$(accept H_0)
Reject H_0	Type I $= P_{H_0}$(reject H_0)	no error

- power $=$ 1-Type II error $= P_{H_1}(\text{reject } H_0)$
- Tradeoff between Type I and Type II errors

Optimality of tests

- A level-α test is any test such that

  $P_{H_0}(\text{reject } H_0) \leq \alpha$

- The mathematical framework developed by Neyman and Pearson allows us to identify **optimal** level-α tests in certain scenarios, meaning that the power of the test is as high as possible.

Simple and composite hypothesis

- Usually the hypothesis H_0 is a subset of a larger statistical model $\mathcal{M}$.
- The complement of H_0 within $\mathcal{M}$ is called the alternative hypothesis. i.e., $\mathcal{M} = H_0 \cup H_1$, and $H_0 \cap H_1 = \emptyset$.
- A hypothesis containing only one distribution is called simple; if it contains more than one distribution it is called composite.
- We will consider 3 cases:

simple H_0 vs simple H_1 $\longleftarrow$ optimal tests can be easily found

simple H_0 vs. composite H_1

composite H_0 vs. composite H_1

optimal tests exist only in special cases.

Simple vs Simple: the NP Lemma

Likelihood ratio statistic and critical region:

$\underset{\sim}{x} = (x_1, \cdots, x_n)$

$H_0: \underset{\sim}{X} \sim f_0(\cdot) \quad H_1: \underset{\sim}{X} \sim f_1(\cdot)$

$f_0(\cdot)$, $f_1(\cdot)$ are pdf (or pmf for discrete distributions)

$$Y = \frac{f_1(\underset{\sim}{X})}{f_0(\underset{\sim}{X})}$$

larger this ratio, more likely f_1 is the underlying distribution.

Critical region has the form:

$$C = \left\{ \underset{\sim}{x} : \frac{f_1(\underset{\sim}{x})}{f_0(\underset{\sim}{x})} \geq y \right\}$$

y some value

Simple vs Simple: the NP Lemma

The Neyman-Pearson (NP) Lemma: Let H_0 and H_1 be simple hypotheses (in which the distributions are either both discrete or both continuous). Fix a level $0 < \alpha < 1$, and suppose there exists y_α such that

$$P_{f_0}(Y \geq y_\alpha) = \alpha.$$

Then for any other test of H_0 with significance level at most α, its power against H_1 is at most the power of this likelihood ratio test.

$$C = \left\{ \underset{\sim}{x} : Y = \frac{f_1(\underset{\sim}{x})}{f_0(\underset{\sim}{x})} \geq y_\alpha \right\}$$

$$P_{f_0}(C) = P_{f_0}(\underset{\sim}{X} \in C) = P_{f_0}(Y \geq y_\alpha)$$

Let D be any other set s.t. $P_{f_0}(D) \leq \alpha$

Simple vs Simple: the NP Lemma

Want to prove Power of C $= P_{f_1}(C) \geq$ power of D

$= P_{f_1}(D)$

Pf of NP lemma (continuous case, discrete case similar)

$\underline{\alpha = P_{f_0}(C)} = \int \cdots \int_{\underset{\sim}{x} \in C} f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$ $\qquad \underset{\sim}{x} = (x_1, \cdots, x_n)$

$= \int \cdots \int_{C \cap D} f_0(\underset{\sim}{x})\, d\underset{\sim}{x} + \int \cdots \int_{C \cap D^c} f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$ ①

$\underline{\alpha \geq P_{f_0}(D)} = \int \cdots \int_{\underset{\sim}{x} \in D} f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$

$= \int \cdots \int_{D \cap C} f_0(\underset{\sim}{x})\, d\underset{\sim}{x} + \int \cdots \int_{D \cap C^c} f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$ ②

① − ② $\quad P_{f_0}(C) - P_{f_0}(D) = \int \cdots \int_{C \cap D^c} f_0(\underset{\sim}{x})\, d\underset{\sim}{x} - \int \cdots \int_{D \cap C^c} f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$

≥ 0 ③

In the same way,

$\underline{P_{f_1}(C) - P_{f_1}(D)} = \int \cdots \int_{C \cap D^c} f_1(\underset{\sim}{x})\, d\underset{\sim}{x} - \int \cdots \int_{D \cap C^c} f_1(\underset{\sim}{x})\, d\underset{\sim}{x}$

$= \int \cdots \int_{C \cap D^c} \underbrace{\frac{f_1(\underset{\sim}{x})}{f_0(\underset{\sim}{x})}}_{\geq y_\alpha} \cdot f_0(\underset{\sim}{x})\, d\underset{\sim}{x} - \int \cdots \int_{D \cap C^c} \underbrace{\frac{f_1(\underset{\sim}{x})}{f_0(\underset{\sim}{x})}}_{< y_\alpha} \cdot f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$

$\geq y_\alpha \int \cdots \int_{C \cap D^c} f_0(\underset{\sim}{x})\, d\underset{\sim}{x} - y_\alpha \int \cdots \int_{D \cap C^c} f_0(\underset{\sim}{x})\, d\underset{\sim}{x}$

$= y_\alpha \left(P_{f_0}(C) - P_{f_0}(D) \right)$ using ③

≥ 0

using 3 again

Y_α nonnegative since

likelihood ratio nonnegative.

$//$

Example 1

$X_1, \ldots, X_n$ iid RVs. $H_0 : X_i \sim N(0,1)$. $H_1 : X_i \sim N(\mu_0, 1)$ for some fixed $\mu_0 \neq 0$.

$$f_0(\underset{\sim}{x}) = \prod_{i=1}^{n} \phi(x_i) \qquad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

$$f_1(\underset{\sim}{x}) = \prod_{i=1}^{n} \phi(x_i - \mu_0)$$

$$Y = \frac{f_1(\underset{\sim}{x})}{f_0(\underset{\sim}{x})} = \prod_{i=1}^{n} \frac{\phi(x_i - \mu_0)}{\phi(x_i)}$$

$$= \frac{e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu_0)^2}}{e^{-\frac{1}{2}\sum_{i=1}^{n} x_i^2}}$$

$$= \exp\left(\mu_0 \sum_{i=1}^{n} x_i - \frac{n\mu_0^2}{2}\right)$$

Reject H_0 if $Y \geq y_\alpha$

with $P_0(Y \geq y_\alpha) = \alpha$, $\qquad P_0(\cdot) = P_{f_0}(\cdot)$

$$P_0\left(\exp\left(\mu_0 \sum_{i=1}^{n} X_i - \frac{n\mu_0^2}{2}\right) \geq y_\alpha\right)$$

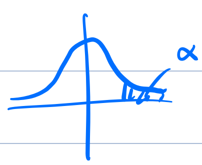
$$= P_0\left(\mu_0 \sum_{i=1}^{n} X_i \geq \log y_\alpha + \frac{n\mu_0^2}{2}\right) \quad (*)$$

Case 1 $\quad \mu_0 > 0 \quad$ let $T = \sum_{i=1}^{n} X_i$.

$$(*) = P_0\left(T \geq \underbrace{\frac{\log y_\alpha + \frac{n\mu_0^2}{2}}{\mu_0}}_{t_\alpha^+}\right)$$

$$= P_0\left(Z \geq \frac{t_\alpha^+}{\sqrt{n}}\right) = \alpha,$$

$T \sim N(0, n)$ under H_0



$\frac{t_\alpha^+}{\sqrt{n}}$ = qnorm$(1-\alpha)$ in R

reject H_0 when T is large enough.

Note that the rejection region does not depend on the value of μ_0, only its sign.