Agenda

1) Hypothesis Testing motivation

2) Type 1 and type 2 errors.

3) NP Lemma

4) Problems

Intro to Hypothesis Testing

Up until now we've just been getting estimates and confidence intervals.

How can I test if my estimate is just noise?

We can use hypothesis testing.

We focus on Neyman-Pearson framework, Null Hypothesis Significance Testing

So we formulate a problem of a null hypothesis and an alternative hypothesis.

H_0 (null)

H_1 (alternative)

we're looking to reject null.

2 types of hypotheses

Simple hypothesis:

Completely specifies the distribution when we assume it.

Suppose our model

$X \sim bin(10, p)$

Simple hypothesis $p = 0.5$

Composite hypotheses:

Not simple

in previous example $p \in (0.5, 1)$

When do we pick to reject/not reject the null?

Def.: The set of values of some statistic T for which H_0 is rejected is called a rejection region.

The set of values for which H_0 is not rejected is the acceptance region.

We never say accept the null. We either reject or fail to reject.

In this framework what can go wrong?

We can make one of two mistakes.

1. we reject H_0 when it is actually true

(Type 1 error)

2. we fail to reject H_0 when it is actually false

(Type 2 error)

Which is worse?

Courtroom Analogy

Rejecting the null is conviction

Failing to Reject is acquittal

A Type 1 error is sending an innocent person to jail.

A Type 2 error is letting someone guilty go free.

Notation

$$P_{H_0}(\text{event})$$

"Probability of event under the null"

$P(\text{type I error})$

$= P_{H_0}(\text{rejecting } H_0) = \alpha$

This is called the significance level.

Under composite null

$$\alpha = \max_{H \in H_0} P_H(\text{type I error})$$

$P(\text{type II error})$

$= P_{H_1}(\text{fail to reject } H_0) = \beta$

Typically we work with

$$1-\beta = P_{H_1}(\text{reject } H_0)$$

we want $\alpha, \beta = 0$, this is impossible.

1) Fix a priori α (to small value)

2) $\max(1-\beta)$ (power) subject to our constrained α.

Q3 from Ch9.11

Suppose $X \sim \text{bin}(100, p)$

Consider some test that rejects $H_0: p = 0.5$

$H_1: p \neq 0.5$

for $|x - 50| > 10$.

Use normal approx to X

find α, graph power function.

Sol:

$\alpha = P_{H_0}$ (rejecting H_0)

$$= P_{H_0}\left(|X-50| > 10 \right)$$

$$= 1 - P_{H_0}\left(|X-50| \leq 10 \right)$$

$$= 1 - P_{H_0}\left(-10 \leq X-50 \leq 10 \right)$$

$$= 1 - P_{H_0}\left(40 \leq X \leq 60 \right)$$

$$E(X) = np \quad , Var(X) = np(1-p)$$

$$X \overset{CLT}{\sim} N(np, np(1-p))$$

$$= 1 - P_{H_0}\left(\frac{40-np}{\sqrt{np(1-p)}} \leq Z \leq \frac{60-np}{\sqrt{np(1-p)}} \right)$$

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \stackrel{\cdot}{\sim} N(0,1)$$

$$= 1 - P\left(\frac{40-50}{5} \leq Z \leq \frac{60-50}{5}\right)$$

$$= 1 - P(-2 \leq Z \leq 2)$$

$$\approx 1 - \left(\Phi(2) - \Phi(-2)\right)$$

Standard normal CDF

$$= 0.0455$$

Now power function

$$1-\beta = P_{H_1}(\text{reject } H_0)$$

$$= 1 - P_{H_1}\left(\frac{40-np}{\sqrt{np(1-p)}} \le Z \le \frac{60-np}{\sqrt{np(1-p)}}\right)$$

$$= 1 - P_{H_1}\left(\frac{40-100p}{10\sqrt{p(1-p)}} \le Z \le \frac{60-100p}{10\sqrt{p(1-p)}}\right)$$

H_1: $p \neq 0.5$.

$$= 1 - \left(\Phi\left(\frac{40-100p}{10\sqrt{p(1-p)}}\right) - \Phi\left(\frac{60-100p}{10\sqrt{p(1-p)}}\right)\right)$$

How do we actually construct a test

Neyman-Pearson Lemma.
gets us most powerful tests for simple-simple hypothesis pairs.

Ex Simple-Simple Test.

$X \sim \text{binom}(100, p)$

$H_0: p = 0.5 \quad H_1: p = 0.7$

Simple-Composite Test

$H_0: p = 0.5 \quad H_1: p \neq 0.5$

Likelihood Ratio Test

Suppose simple-simple hypothesis.

We can get a likelihood under both alternative and null

Denote $f_{H_0}(\underline{X})$ or $f_0(\underline{X})$ as the likelihood under H_0.

$$LR(\underline{X}) = \frac{f_{H_0}(\underline{X})}{f_{H_1}(\underline{X})}$$

if H_1 is more "likely"

$$f_{H_1}(x) > f_{H_0}(x)$$

equiv

$$\frac{f_{H_0}(x)}{f_{H_1}(x)} \text{ is small}$$

NP-Lemma

Suppose that we have a test that rejects H_0 when

$$\frac{f_{H_0}(x)}{f_{H_1}(x)} < c(\alpha)$$

has significance level α.

Then this the most powerful α level test for simple-simple hypotheses.

Process

1) Compute Likelihood Ratio

2) Find statistic $T(x)$ that determines when LR is small

3) Determine rejection region for T

4) Choose cutoff $c(\alpha)$ to ensure α level test.

Q7 Ch 9.11

Let $X_1, \dots, X_n$ be a sample from Poisson dist.

Find LRT for

$H_0: \lambda = \lambda_0$. $H_1: \lambda = \lambda_1$

know $\lambda_1 > \lambda_0$.

if $X_i \overset{ind}{\sim} pois(\lambda)$

$$P(X_i = x_i) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$L_n(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$= \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod x_i !}$$

$$LR(x) = \frac{f_{H_0}(x)}{f_{H_1}(x)} = \frac{L_n(\lambda_0)}{L_n(\lambda_1)}$$

$$= \frac{e^{-n\lambda_0} \lambda_0^{\sum x_i} / \cancel{\prod x_i !}}{e^{-n\lambda_1} \lambda_1^{\sum x_i} / \cancel{\prod x_i !}}$$

$$= e^{-n\lambda_0 - n\lambda_1} \cdot \left(\frac{\lambda_0}{\lambda_1}\right)^{\sum x_i} \quad (1)$$

Let $T = \sum X_i$

$$= e^{-n\lambda_0 - n\lambda_1} \left(\frac{\lambda_0}{\lambda_1}\right)^T$$

$$\lambda_1 > \lambda_0 \Rightarrow \frac{\lambda_0}{\lambda_1} < 1 \qquad (2)$$

Then $LR(x)$ is small when T is large

LRT rejects when

$T(x) = \sum X_i$ is large (3)

or

$$\sum X_i > c(\alpha)$$

(d) $P_{H_0}(\text{rejecting } H_0) = \alpha$

$$= P_{H_0}\left(\sum_{i=1}^{n} X_i \geq c(\alpha)\right)$$

$$T = \sum_{i=1}^{n} X_i \sim \text{pois}(n\lambda_0)$$

under the null

$$= P_{H_0}(T \geq c(\alpha))$$

$$= 1 - P_{H_0}(T \leq c(\alpha)) = \alpha.$$

$$= 1 - F(c(\alpha)) = \alpha.$$

we can just solve for $c(\alpha)$

given any α.

$$= 1 - \sum_{t=0}^{c(\alpha)} \frac{(n\lambda_0)^t e^{-n\lambda_0}}{t!} = \alpha$$

Power

P_{H_1} (reject H_0)

now known

$$= P_{H_1}(T > c(\alpha))$$

Under alternative

$$T \sim pois(n\lambda_1)$$

$$= 1 - P_{H_1}(T \le c(\alpha))$$

$$= 1 - F(c(\alpha))$$

$$= 1 - \sum_{t=0}^{c(\alpha)} e^{-nd_1}(nd_1)^t / t!$$

Dog Cancer problem

Want to test whether or not dogs smelled cancer due to chance.

$$X_1, \dots, X_{54} \overset{ind}{\sim} \text{Bern}(p)$$

$$H_0: p = \frac{1}{7} \quad H_1: p > \frac{1}{7}$$

Use LRT even
though not simple.

$H_0: p = p_0.$

$H_1: p = p_1 \quad , \quad p_1 > p_0.$

$$L_n(p; \underset{\sim}{x}) = \prod_{i=1}^{54} p^{x_i} (1-p)^{1-x_i}$$

$$= p^{\Sigma x_i} (1-p)^{54 - \Sigma x_i}$$

$$LR(\underset{\sim}{x}) = \frac{f_{H_0}(\underset{\sim}{x})}{f_{H_1}(\underset{\sim}{x})} = \frac{p_0^{\Sigma x_i} (1-p_0)^{54-\Sigma x_i}}{p_1^{\Sigma x_i} (1-p_1)^{54-\Sigma x_i}}$$

$$= \left(\frac{p_0}{p_1}\right)^{\Sigma x_i} \left(\frac{1-p_0}{1-p_1}\right)^{54-\Sigma x_i}$$

$$= \left(\frac{p_1}{p_0}\right)^{-\Sigma x_i} \left(\frac{1-p_0}{1-p_1}\right)^{54-\Sigma x_i}$$

$$= \left(\frac{p_1}{p_0}\right)^{-54} \left(\frac{p_1}{p_0}\right)^{54-\Sigma x_i} \left(\frac{1-p_0}{1-p_1}\right)^{54-\Sigma x_i}$$

$$= \left(\frac{p_1}{p_0}\right)^{-54} \left(\frac{p_1}{p_0} \frac{1-p_0}{1-p_1}\right)^{54-\Sigma x_i}$$

By assumption

$p_1 > p_0 \Rightarrow \frac{p_1}{p_0} > 1$

$\frac{1-p_0}{1-p_1} > 1$

$\Rightarrow$ dec function of $\sum x_i$

So we reject for small

$LR(\underline{x})$

$\Longleftrightarrow$

reject for large values

of $\sum x_i = T(\underline{x})$

$$P_{H_0}\left(T > c(\alpha)\right) = \alpha$$

$$= 1 - P_{H_0}\left(T \leq c(\alpha)\right) = \alpha$$

under null

$$T = \sum X_i \sim \text{binom}\left(54, \frac{1}{7}\right)$$

$$= 1 - F(c(\alpha)) = \alpha$$

$$1 - \sum_{k=0}^{c(\alpha)} \binom{n}{k} \left(\frac{1}{7}\right)^k \left(\frac{6}{7}\right)^{54-k} = \alpha$$

Ch 9.11 Q6.c), d)

There are two coins

one has $P(\text{heads}=0.5)$

other has $P(\text{heads}=0.7)$

we test

$H_0: p = 0.5$

$H_1: p = 0.7$

We flip a coin until a head comes up.

Call the number of flips X.

$X \sim \text{geom}(p)$

Given a test which rejects if $X \geq 8$

what is the significance level

$$P_{H_0}(\text{rejecting})$$

$$= P_{H_0}(X \geq 8)$$

$$= \sum_{x=8}^{\infty} p\,(1-p)^{x-1}$$

(under null $p=0.5$)

$$= \sum_{x=8}^{\infty} 0.5^{x-1} \, 0.5$$

$$= \sum_{x=8}^{\infty} (0.5)^{x}$$

$$= (0.5)^{8} \sum_{x=8}^{\infty} (0.5)^{x-8}$$

$$= (0.5)^{8} \sum_{x=0}^{\infty} (0.5)^{x}$$

$$= (0.5)^{8} \frac{1}{1-0.5} = (0.5)^{7}$$

$$= 0.00781$$

Now power

$$P_{H_1}(X \geq 8)$$

$$= \sum_{x=8}^{\infty} (0.7)\, 0.3^{x-1}$$

$$= 0.7 \sum_{x=8}^{\infty} 0.3^{x-1}$$

$$= 0.7\,(0.3)^7 \sum_{x=8}^{\infty} 0.3^{x-8}$$

$$= 0.7\,(0.3)^7 \frac{1}{1-0.3}$$

$$= (0.3)^7 = 0.0002187$$