Mathematical Statistics

Anna Janicka

Lecture XI, 6.05.2019

HYPOTHESIS TESTING III:

LR TEST FOR COMPOSITE HYPOTHESES

EXAMPLES OF ONE-SAMPLE TESTS

Plan for today

1. LR test for composite hypotheses
2. Examples of LR tests:
   - Model I: One- and two-sided tests for the mean in the normal model, σ^2 known
   - Model II: One- and two-sided tests for the mean in the normal model, σ^2 unknown

     \+ One- and two-sided tests for the variance
   - Model III: Tests for the mean, large samples
   - Model IV: Tests for the fraction, large samples
3. Asymptotic properties of the LR test
4. Test randomization

Testing simple hypotheses – reminder

We observe X. We want to test

H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$.

(two simple hypotheses)

We can write it as:

H_0: $X \sim f_0$ against H_1: $X \sim f_1$,

where f_0 and f_1 are *densities* of distributions defined by θ_0 and θ_1 (i.e. P_0 and P_1)

Likelihood ratio test for simple hypotheses. Neyman-Pearson Lemma – reminder

H_0: $X \sim f_0$ against H_1: $X \sim f_1$

Let

$$K^* = \left\{ x \in \mathrm{X} : \frac{f_1(x)}{f_0(x)} > c \right\}$$

such that $P_0(K^*) = \alpha$ and $P_1(K^*) = 1 - \beta$

Then, for any $K \subseteq \mathrm{X}$:

if $P_0(K) \leq \alpha$, then $P_1(K) \leq 1 - \beta$.

(i.e.: the test with critical region K^* is the most powerful test for testing H_0 against H_1)

Likelihood ratio test for composite hypotheses

$X \sim P_\theta$, $\{P_\theta: \theta \in \Theta\}$ – family of distributions

We are testing H_0: $\theta \in \Theta_0$ against H_1: $\theta \in \Theta_1$

such that $\Theta_0 \cap \Theta_1 = \emptyset$, $\Theta_0 \cup \Theta_1 = \Theta$

Let

H_0: $X \sim f_0(\theta_0,\cdot)$ for some $\theta_0 \in \Theta_0$.

H_1: $X \sim f_1(\theta_1, \cdot)$ for some $\theta_1 \in \Theta_1$,

where f_0 and f_1 are densities (for $\theta \in \Theta_0$ and $\theta \in \Theta_1$, respectively)

Just like in the N-P Lemma, but models are statistic – contain unknow parameters. We proceed similarly.

Neyman-Pearson Lemma – Example 1 – reminder

Normal model: $X_1, X_2, \ldots, X_n$ are an IID sample from $N(\mu, \sigma^2)$, σ^2 is known

The most powerful test for

$$H_0\colon \mu = 0 \text{ against } H_1\colon \mu = 1.$$

$\mu_0 < \mu_1$

At significance level α :

$$K^* = \left\{ (x_1, x_2, \ldots, x_n) : \overline{X} > {u_{1-\alpha}\sigma}/{\sqrt{n}} \right\}$$

For obs. 1.37; 0.21; 0.33; -0.45; 1.33; 0.85; 1.78; 1.21; 0.72 from $N(\mu, 1)$ we have, for $\alpha = 0.05$:

$$\overline{X} \approx 0.82 > {1.645 \cdot 1}/{\sqrt{9}} \approx 0.54$$

→ we reject H_0

Model I

Neyman-Pearson Lemma – Example 1 cont. – reminder

Power of the test

$$P_1(K^*) = P\left(\overline{X} > {}^{1.645\sigma}\!\Big/\!_{\sqrt{n}} \,\Big|\, \mu = 1 \right) = ...$$

$$= 1 - \Phi\left(1.645 - {}^{\mu_1 \cdot \sqrt{n}}\!\Big/\!_{\sigma} \right) \qquad \approx 0.91$$

If we change α, μ_1, n – the power of the test....

Neyman-Pearson Lemma: Generalization of example 1

The same test is UMP for H_1: $\mu > 0$ and for

$$H_0: \mu \leq 0 \text{ against } H_1: \mu > 0$$

more generally: under additional assumptions about the family of distributions, the same test is UMP for testing

$$H_0: \mu \leq \mu_0 \text{ against } H_1: \mu > \mu_0$$

Note the change of direction in the inequality when testing

$$H_0: \mu \geq \mu_0 \text{ against } H_1: \mu < \mu_0$$

Neyman-Pearson Lemma – Example 2

Exponential model: $X_1, X_2, \ldots, X_n$ are an IID sample from distr exp(λ), $n = 10$.

MP test for

H_0: $\lambda = ½$ against H_1: $\lambda = ¼$.

At significance level $\alpha = 0.05$:

$$K^* = \left\{ (x_1, x_2, \ldots, x_{10}) : \sum x_i > 31.41 \right\}$$

E.g. for a sample: 2; 0.9; 1.7; 3.5; 1.9; 2.1; 3.7; 2.5; 3.4; 2.8:

$\Sigma = 24.5 \rightarrow$ no grounds for rejecting H_0.

$$\exp(\lambda) = \Gamma(1, \lambda) \qquad \Gamma(a, \lambda) + \Gamma(b, \lambda) = \Gamma(a + b, \lambda) \qquad \Gamma(n/2, 1/2) = \chi^2(n)$$

Neyman-Pearson Lemma – Example 2'

Exponential model: $X_1, X_2, \ldots, X_n$ are an IID sample from distr exp(λ), $n = 10$.

MP test for

$$H_0: \lambda = ½ \text{ against } H_1: \lambda = ¾.$$

At significance level $\alpha = 0.05$:

$$K^* = \left\{(x_1, x_2, \ldots, x_{10}) : \sum x_i < 10.85\right\}$$

E.g. for a sample: 2; 0.9; 1.7; 3.5; 1.9; 2.1; 3.7; 2.5; 3.4; 2.8:

$\Sigma = 24.5 \rightarrow$ no grounds for rejecting H_0.

$$\exp(\lambda) = \Gamma(1, \lambda) \qquad \Gamma(a, \lambda) + \Gamma(b, \lambda) = \Gamma(a + b, \lambda) \qquad \Gamma(n/2, 1/2) = \chi^2(n)$$

Example 2 cont.

The test

$$K^* = \left\{(x_1, x_2, \ldots, x_{10}) : \sum x_i > 31.41\right\}$$

is UMP for H_0: $\lambda \geq$ ½ against H_1: $\lambda <$ ½

The test

$$K^* = \left\{(x_1, x_2, \ldots, x_{10}) : \sum x_i < 10.85\right\}$$

is UMP for H_0: $\lambda \leq$ ½ against H_1: $\lambda >$ ½

Likelihood ratio test for composite hypotheses – cont.

Test statistic: $$\lambda = \frac{\sup_{\theta_1 \in \Theta_1} f_1(\theta_1, X)}{\sup_{\theta_0 \in \Theta_0} f_0(\theta_0, X)}$$

or $$\lambda = \frac{f_1(\hat{\theta}_1, X)}{f_0(\hat{\theta}_0, X)}$$

where $\hat{\theta}_0, \hat{\theta}_1$ are MLE of the null and alternative hypothesis models

We reject H_0 if $\lambda > c$ for a constant c (determined according to significance level)

Likelihood ratio test for composite hypotheses – justification

Just like in the Neyman-Pearson Lemma, we compare the "highest chance of obtaining observation X, when the alternative is true" to the "highest chance of obtaining observation X, when the null is true"; we reject the null hypothesis in favor of the alternative if this ratio is very unfavorable for the null.

Likelihood ratio test for composite hypotheses – alternative version

Test statistic: $$\tilde{\lambda} = \frac{\sup_{\theta \in \Theta} f(\theta, X)}{\sup_{\theta_0 \in \Theta_0} f_0(\theta_0, X)}$$

or $$\tilde{\lambda} = \frac{f(\hat{\theta}, X)}{f_0(\hat{\theta}_0, X)}$$

where $\hat{\theta}, \hat{\theta}_0$ are the ML estimators for the model without restrictions and for the null model.

We reject H_0 if $\tilde{\lambda} > \tilde{c}$ for a constant $\tilde{c}$.

more convenient if the null is simple or if models are nested

Likelihood ratio test for composite hypotheses – properties

For some models with composite hypotheses the UMPT *does not exist* (so the LR test will not be UMP because there is no such test)

e.g. testing H_0: $\theta = \theta_0$ against H_1: $\theta \neq \theta_0$ if the family of distributions has a *monotonic LR property*, i.e. $f_1(x)/f_0(x)$ is an increasing function of a statistic $T(x)$ for any f_0 and f_1 corresponding to parameters $\theta_0 < \theta_1$.

In order to have UMPT for H_0: $\theta = \theta_0$ against H_1: $\theta < \theta_0$ we would need a critical region of the type $T(x)>c$, and to have a UMPT for H_0: $\theta = \theta_0$ against H_1: $\theta > \theta_0$ we would need a critical region of the type $T(x)<c$, so it is impossible to find a UMPT for H_1: $\theta \neq \theta_0$.

Likelihood ratio test: special cases

The exact form of the test depends on the distribution.

In many cases, finding the distribution is hard/complicated (in many such cases, we use the asymptotic properties of the LR test instead of precise formulae)

Notation

$x_{something}$ **always** means a quantile of rank something

Model I: comparing the mean

Normal model: $X_1, X_2, ..., X_n$ are an IID sample from $N(\mu, \sigma^2)$, where σ^2 is **known**

$H_0: \mu = \mu_0$

Test statistic:

$$U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \sim N(0,1)$$

$H_0: \mu = \mu_0$ against $H_1: \mu > \mu_0$

critical region $K^* = \{x : U(x) > u_{1-\alpha}\}$

$H_0: \mu = \mu_0$ against $H_1: \mu < \mu_0$

critical region $K^* = \{x : U(x) < u_{\alpha} = -u_{1-\alpha}\}$

$H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$

critical region $K^* = \{x : |U(x)| > u_{1-\alpha/2}\}$

Example 1

Model I: example

Let $X_1, X_2, \ldots, X_{10}$ be an IID sample from $N(\mu, 1^2)$:

-1.21 -1.37 0.51 0.37 -0.75 0.44 1.20 -0.96 -1.14 -1.40

Is $\mu = 0$? (for $\alpha = 0.05$)

In the sample: mean = -0.43, ~~variance = 0.92~~

Test statistic:

$$U = \frac{-0.43 - 0}{1}\sqrt{10} \approx -1.36$$

H_0: $\mu = 0$ against H_1: $\mu \neq 0$, $u_{0.975} \approx 1.96$ (p-value $\approx$ 0.172)

H_0: $\mu = 0$ against H_1: $\mu < 0$, $u_{0.05} \approx -1.64$ (p-value $\approx$ 0.086)

H_0: $\mu = 0$ against H_1: $\mu > 0$, $u_{0.95} \approx 1.64$ (p-value $\approx$ 0.914)

→ in none of these cases are there grounds to reject H_0 for $\alpha = 0.05$

→but we would reject H_0: $\mu = 0$ in favor of H_1: $\mu < 0$ for $\alpha = 0.1$

Model II: comparing the mean

Normal model: $X_1, X_2, ..., X_n$ are an IID sample from $N(\mu, \sigma^2)$, where σ^2 is **unknown**

H_0: $\mu = \mu_0$

Test statistic:

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n} \sim t(n-1)$$

H_0: $\mu = \mu_0$ against H_1: $\mu > \mu_0$

critical region $K^* = \{x : T(x) > t_{1-\alpha}(n-1)\}$

H_0: $\mu = \mu_0$ against H_1: $\mu < \mu_0$

critical region $K^* = \{x : T(x) < t_{\alpha}(n-1)\}$

H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$

critical region $K^* = \{x : |T(x)| > t_{1-\alpha/2}(n-1)\}$

$$t_{\alpha}(n-1) = -t_{1-\alpha}(n-1)$$

Model II: example (mean)

Let $X_1, X_2, ..., X_{10}$ be an IID sample from $N(\mu, \sigma^2)$:

-1.21 -1.37 0.51 0.37 -0.75 0.44 1.20 -0.96 -1.14 -1.40

Is $\mu = 0$? (for $\alpha = 0.05$)

In the sample: mean = -0.43, variance = 0.92

Test statistic:

$$U = \frac{-0.43 - 0}{\sqrt{0.92}}\sqrt{10} \approx -1.42$$

$H_0: \mu = 0$ vs $H_1: \mu \neq 0$, $t_{0.975}(9) \approx 2.26$ (p-value $\approx$ 0.188)

$H_0: \mu = 0$ vs $H_1: \mu < 0$, $t_{0.05}(9) \approx -1.83$ (p-value $\approx$ 0.094)

$H_0: \mu = 0$ vs $H_1: \mu > 0$, $t_{0.95}(9) \approx 1.83$ (p-value $\approx$ 0.906)

→ in none of these cases are there grounds to reject H_0 for $\alpha = 0.05$

→but we would reject H_0: $\mu = 0$ in favor of H_1: $\mu < 0$ for $\alpha = 0.1$

Model II: comparing the variance

Normal model: $X_1, X_2, \ldots, X_n$ are an IID sample from $N(\mu, \sigma^2)$, where σ^2 is **unknown**

H_0: $\sigma = \sigma_0$

Test statistic:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

H_0: $\sigma = \sigma_0$ against H_1: $\sigma > \sigma_0$

critical region $K^* = \{x : \chi^2(x) > \chi^2_{1-\alpha}(n-1)\}$

H_0: $\sigma = \sigma_0$ against H_1: $\sigma < \sigma_0$

critical region $K^* = \{x : \chi^2(x) < \chi^2_{\alpha}(n-1)\}$

H_0: $\sigma = \sigma_0$ against H_1: $\sigma \neq \sigma_0$

critical region $K^* = \{x : \chi^2(x) < \chi^2_{\alpha/2}(n-1) \vee \chi^2(x) > \chi^2_{1-\alpha/2}(n-1)\}$

Model II: example (variance)

Let $X_1, X_2, ..., X_{10}$ be an IID sample from $N(\mu, \sigma^2)$:

-1.21 -1.37 0.51 0.37 -0.75 0.44 1.20 -0.96 -1.14 -1.40

Is $\sigma = 1$? (for $\alpha = 0.05$)

In the sample: variance = 0.92

Test statistic: $$\chi^2 = \frac{9 \cdot 0.92}{1} \approx 8.28$$

H_0: $\sigma = 1$ against H_1: $\sigma > 1$ $\quad \chi^2_{0.95} \approx 16.92$

H_0: $\sigma = 1$ against H_1: $\sigma < 1$ $\quad \chi^2_{0.05} \approx 3.33$

H_0: $\sigma = 1$ against H_1: $\sigma \neq 1$ $\quad \chi^2_{0.025} \approx 2.70; \chi^2_{0.975} \approx 19.02$

→ in none of these cases are there grounds to reject H_0 (for $\alpha = 0.05$)

Model III: comparing the mean

Asymptotic model: $X_1, X_2, \ldots, X_n$ are an IID sample from a distribution with mean μ and variance (unknown), n – large.

H_0: $\mu = \mu_0$

Test statistic:

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n}$$

has, for large *n, an approximate* distribution N(0,1)

H_0: $\mu = \mu_0$ against H_1: $\mu > \mu_0$

critical region $$K^* = \{x : T(x) > u_{1-\alpha}\}$$

H_0: $\mu = \mu_0$ against H_1: $\mu < \mu_0$

critical region $$K^* = \{x : T(x) < u_{\alpha} = -u_{1-\alpha}\}$$

H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$

critical region $$K^* = \{x : |T(x)| > u_{1-\alpha/2}\}$$

Model IV: comparing the fraction

Asymptotic model: $X_1, X_2, \ldots, X_n$ are an IID sample from a two-point distribution, n – large.

$$P_p(X=1)=p=1-P_p(X=0)$$

H_0: $p = p_0$

Test statistic:

$$U^*=\frac{\bar{X}-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}=\frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$$

has *an approximate* distribution N(0,1) for large n

H_0: $p = p_0$ against H_1: $p > p_0$

critical region $K^*=\{x: U(x)>u_{1-\alpha}\}$

H_0: $p = p_0$ against H_1: $p < p_0$

critical region $K^*=\{x: U(x)<u_{\alpha}=-u_{1-\alpha}\}$

H_0: $p = p_0$ against H_1: $p \neq p_0$

critical region $K^*=\{x: |U(x)|>u_{1-\alpha/2}\}$

Model IV: example

We toss a coin 400 times. We get 180 heads. Is the coin symmetric?

H_0: p = ½

$$U^* = \frac{(180/400 - 1/2)}{\sqrt{1/2(1-1/2)}}\sqrt{400} = -2$$

for $\alpha = 0.05$ and H_1: $p \neq$ ½ we have $u_{0.975} = 1.96 \rightarrow$ we reject H_0

for $\alpha = 0.05$ and H_1: $p <$ ½ we have $u_{0.05} = -u_{0.95} = -1.64$

$\rightarrow$ we reject H_0

for $\alpha = 0.01$ and H_1: $p \neq$ ½ we have $u_{0.995} = 2.58$

$\rightarrow$ we do not reject H_0

for $\alpha = 0.01$ and H_1: $p <$ ½ we have $u_{0.01} = -u_{0.99} = -2.33$

$\rightarrow$ we do not reject H_0

p-value for H_1: $p \neq$ ½: 0.044 **p-value for H_1: $p <$ ½: 0.022**

Asymptotic properties of the LR test

We consider two nested models, we test

H_0: $h(\theta) = 0$ against H_1: $h(\theta) \neq 0$

Under the assumption that

- *h* is *a nice function*
- Θ is a *d*-dimensional set
- $\Theta_0 = \{\theta : h(\theta) = 0\}$ is a $d - p$ dimensional set

Theorem: If H_0 is true, then for $n \to \infty$ the distribution of the statistic $2\ln\tilde{\lambda}$ converges to a chi-squared distribution with *p* degrees of freedom

degrees of freedom = number of restrictions

Asymptotic properties of the LR test – example

Exponential model: $X_1, X_2, \ldots, X_n$ are an IID sample from Exp(θ).

We test H_0: $\theta = 1$ against H_1: $\theta \neq 1$

$$MLE(\theta) = \hat{\theta} = 1/\overline{X}$$

$$\tilde{\lambda} = \frac{\Pi f_{\hat{\theta}}(x_i)}{\Pi f_1(x_i)} = \frac{\frac{1}{\overline{X}^n}\exp(-\frac{1}{\overline{X}}\Sigma x_i)}{\exp(-\Sigma x_i)} = \frac{1}{\overline{X}^n}\exp\left(n(\overline{X}-1)\right)$$

then: $\tilde{\lambda} > \tilde{c} \Leftrightarrow 2\ln\tilde{\lambda} > 2\ln\tilde{c}$

from Theorem: $2\ln\tilde{\lambda} = 2n((\overline{X}-1) - \ln\overline{X}) \xrightarrow{D} \chi^2(1)$

for a sign. level α =0.05 we have $\chi^2_{0.95}(1) \approx 3.84 \approx 2\ln\tilde{c}$

so we reject H_0 in favor of H_1 if $\tilde{\lambda} > e^{3.84/2}$

Test randomization

Sometimes, a test with a significance level exactly equal to α does not exist (e.g. for discrete random variables).

In such cases, we need *randomization*.

(The UMPT, if it exists, needs to be randomized).

eg. no of heads in 8 tosses, $H_0 : p = ½$, $H_1 : p < ½$, α=0.05:

x_i	0	1	2	3	4	5	6	7	8
p_i	0.004	0.03	0.11	0.22	0.27	0.22	0.11	0.03	0.004
cum p_i	0.004	0.04	0.15	0.36	0.64	0.86	0.97	0.996	1.000

$X \leq 1$ reject, $X > 2$ OK, $X = 2$: $p = 1/11$ reject