ACTL2131 2.5 - Alternate Hypothesis Testing

Tadhg Xu-Glassop

2025T1

End of term stuff

Please do the MyExperience!

I have online consultations 2-4pm on Monday.

Other Hypothesis Tests

Two Sample Test of Means

A regular hypothesis test using the two sample t stat from the formulae

Two Sample Test of Variances

A regular hypothesis test using two sample variance F stat from the formulae

Other Hypothesis Tests (cont.)

Wald Test

Use asymptotic normality property of MLE and stantardise to "pivot".

For example, if we estimate a parameter with MLE, then $\hat{\theta} \sim \mathcal{N}(\theta, \mathrm{CRLB}(\theta))$. So we use the (approximate) statistic

$$\frac{\hat{\theta} - \theta}{\sqrt{\mathrm{CRLB}(\hat{\theta})}} \sim \mathcal{N}(0, 1).$$

Motivating GLRT's

We have so many hypothesis tests for so many things! And we have to do them so many times and consider all these different conditions and assumptions!

Further, what if we fit a model (a distribution) to a dataset with some sample. How can we say that a set of parameters Ω_0 is worse than another set Ω_1 ? What if we just got unlucky and Ω_1 happened to fit better to the sample, but really Ω_0 is the best?

Motivating GLRT's

We have so many hypothesis tests for so many things! And we have to do them so many times and consider all these different conditions and assumptions!

Further, what if we fit a model (a distribution) to a dataset with some sample. How can we say that a set of parameters Ω_0 is worse than another set Ω_1 ? What if we just got unlucky and Ω_1 happened to fit better to the sample, but really Ω_0 is the best?

We could test multiple parameters at the same time AND compare the goodness-of-fit of different sets of parameters if only we had some way of saying how **likely** a set of parameters are...

Revision: Likelihood

Suppose I have a vector of iid observations (a sample) $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, from a distribution with PDF $f(x; \tilde{\theta})$, where $\tilde{\theta}$ is a vector of unknown parameters. Then, the likelihood of my parameters given my sample is

$$\begin{aligned}
\mathcal{L}(\tilde{\theta}) &= \mathbb{P}[\mathbf{\Theta} = \tilde{\theta} | \mathbf{X} = \mathbf{x}] \\
&= \prod_{i=1}^{n} \mathbb{P}[X_i = x_i | \mathbf{\Theta} = \tilde{\theta}] \\
&= \prod_{i=1}^{n} f_X(x_i; \tilde{\theta}).
\end{aligned}$$

Revision: Likelihood

Suppose I have a vector of iid observations (a sample) $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, from a distribution with PDF $f(x; \tilde{\theta})$, where $\tilde{\theta}$ is a vector of unknown parameters. Then, the likelihood of my parameters given my sample is

$$\begin{aligned}
\mathcal{L}(\tilde{\theta}) &= \mathbb{P}[\boldsymbol{\Theta} = \tilde{\theta} | \mathbf{X} = \mathbf{x}] \\
&= \prod_{i=1}^{n} \mathbb{P}[X_i = x_i | \boldsymbol{\Theta} = \tilde{\theta}] \\
&= \prod_{i=1}^{n} f_X(x_i; \tilde{\theta}).
\end{aligned}$$

- Let $\tilde{\theta}_1$ and $\tilde{\theta}_2$ be two different vectors of parameters. If $\mathcal{L}(\tilde{\theta}_1) \gg \mathcal{L}(\tilde{\theta}_2)$, what can you say?

Revision: Likelihood

Suppose I have a vector of iid observations (a sample) $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, from a distribution with PDF $f(x;\tilde{\theta})$, where $\tilde{\theta}$ is a vector of unknown parameters. Then, the likelihood of my parameters given my sample is

$$\begin{aligned}\mathcal{L}(\tilde{\theta}) &= \mathbb{P}[\boldsymbol{\Theta} = \tilde{\theta} | \mathbf{X} = \mathbf{x}] \\ &= \prod_{i=1}^{n} \mathbb{P}[X_i = x_i | \boldsymbol{\Theta} = \tilde{\theta}] \\ &= \prod_{i=1}^{n} f_X(x_i;\tilde{\theta}).\end{aligned}$$

- Let $\tilde{\theta}_1$ and $\tilde{\theta}_2$ be two different vectors of parameters. If $\mathcal{L}(\tilde{\theta}_1) \gg \mathcal{L}(\tilde{\theta}_2)$, what can you say?
- What is the range of $\mathcal{L}(\tilde{\theta})$?

Comparing Different Likelihoods

We need a nice way of comparing different likelihoods...

Likelihood Ratio Test Statistic

Let Ω_0, Ω_1 be sets of parameters and $\Omega = \Omega_0 \cup \Omega_1$. Then, the likelihood ratio test statistic is

$$\Lambda = \frac{\max_{\Theta \in \Omega_0} \mathcal{L}(\Theta)}{\max_{\Theta \in \Omega} \mathcal{L}(\Theta)}.$$

To maximise the likelihood under Ω, we use maximum likelihood estimation and restrict possible values to those in Ω.

Comparing Different Likelihoods

We need a nice way of comparing different likelihoods...

Likelihood Ratio Test Statistic

Let Ω_0, Ω_1 be sets of parameters and $\Omega = \Omega_0 \cup \Omega_1$. Then, the likelihood ratio test statistic is

$$\Lambda = \frac{\max_{\Theta \in \Omega_0} \mathcal{L}(\Theta)}{\max_{\Theta \in \Omega} \mathcal{L}(\Theta)}.$$

To maximise the likelihood under Ω, we use maximum likelihood estimation and restrict possible values to those in Ω.

- Is the numerator or denominator larger? What is the range of Λ?

Comparing Different Likelihoods

We need a nice way of comparing different likelihoods...

Likelihood Ratio Test Statistic

Let Ω_0, Ω_1 be sets of parameters and $\Omega = \Omega_0 \cup \Omega_1$. Then, the likelihood ratio test statistic is

$$\Lambda = \frac{\max_{\Theta \in \Omega_0} \mathcal{L}(\Theta)}{\max_{\Theta \in \Omega} \mathcal{L}(\Theta)}.$$

To maximise the likelihood under Ω, we use maximum likelihood estimation and restrict possible values to those in Ω.

- Is the numerator or denominator larger? What is the range of Λ?
- If Ω_0 contains really powerful parameters, what will Λ be? What if Ω_0 contains awful parameters?

Generalised Likelihood Ratio Test

Let Ω_0, Ω_1 be sets of parameters and $\Omega = \Omega_0 \cup \Omega_1$. We aim to test the hypothesis,

$$\mathcal{H}_0 : \Theta \in \Omega_0 \qquad \text{v.s.} \qquad \mathcal{H}_1 : \Theta \in \Omega_1.$$

To test, we use the statistic

$$\Lambda = \frac{\max_{\Theta \in \Omega_0} \mathcal{L}(\Theta)}{\max_{\Theta \in \Omega} \mathcal{L}(\Theta)}.$$

To find significance,

- Find exact distribution (hard!);
- Take asymptotic approximation.

Asymptotic Distribution

Asymptotic Distribution of Λ

Let Ω_0, Ω_1 be sets of parameters and $\Omega = \Omega_0 \cup \Omega_1$. Suppose there are p free parameters under Ω_0 and $p + q$ under Ω. Consider Λ as before. Then, when we have sufficient sample size,

$$-2\log(\Lambda) \sim \chi^2_q.$$

A parameter is "free" when we don't restrict it to be a certain value.

Asymptotic Distribution

Asymptotic Distribution of Λ

Let Ω_0, Ω_1 be sets of parameters and $\Omega = \Omega_0 \cup \Omega_1$. Suppose there are p free parameters under Ω_0 and $p + q$ under Ω. Consider Λ as before. Then, when we have sufficient sample size,

$$-2\log(\Lambda) \sim \chi^2_q.$$

A parameter is "free" when we don't restrict it to be a certain value.

Recall that Λ is closer to 0 when the null is false. So, what values of $-2\log\Lambda$ will we reject our null under?

Contingency Tables

A contingency table summarises categorical outcomes of a sample.
It summarises how many 'types' and the totals were surveyed.

	Disease	No Disease	Total
Smoker	40	60	100
Non-Smoker	30	170	200
Total	70	230	300

Expected vs Observed

What if we have some hypothesis on how the data in our table should be? That is, what if we have some null hypothesis on the relationship between our variables that gives us expected entries...

Expected vs Observed

What if we have some hypothesis on how the data in our table should be? That is, what if we have some null hypothesis on the relationship between our variables that gives us expected entries...

Intuitively, we should reject our null if the observed values differ too much from our expectation. So, we need

- A statistic that measures how much the observed values differ from expectation;
- What constitutes as "too much"

Expected vs Observed (Test for Independence).

$\mathcal{H}_0$: Smoking does not lead to higher disease probability.

Table 1: Observed

	Disease	No Disease	Total
Smoker	34	66	100
Non-Smoker	36	164	200
Total	70	230	300

Table 2: Expected (Under $\mathcal{H}_0$)

	Disease	No Disease	Total
Smoker	23	77	100
Non-Smoker	47	153	200
Total	70	230	300

χ^2 Goodness of Fit Tests

χ^2 Statistic

Consider some 'Expected' (from a null hypothesis) and 'Observed' contingency tables with r rows and c columns, and let E_{ij} and O_{ij} be the entry in the ith row and jth column for each table, respectively. Then, the test statistic

$$T = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)\cdot(c-1)},$$

under the null.

We find the critical value of χ^2 with some significance level α.

Do we reject for big or small values of T?

χ^2 Goodness of Fit Tests (cont.)

We can apply this test to singular columns as well! This is useful for goodness of fit tests, when we see compare how many observations should fall into intervals (under null) v.s. what we observed. Note we need $E_i > 5$ for all i.

Table 3: Observed v.s. Expected with $\mathcal{H}_0$: Observations are $\mathcal{N}(0, 1)$

Category	Observed	Expected
(-∞, -1)	50	19
(-1,0)	30	41
(0,1)	20	41
(1, ∞)	20	19
Total	120	120

We have $\displaystyle\sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \sim \chi_p^2$, where $p =$

χ^2 Goodness of Fit Tests (cont.)

We can apply this test to singular columns as well! This is useful for goodness of fit tests, when we see compare how many observations should fall into intervals (under null) v.s. what we observed. Note we need $E_i > 5$ for all i.

Table 3: Observed v.s. Expected with $\mathcal{H}_0$: Observations are $\mathcal{N}(0, 1)$

Category	Observed	Expected
(-$\infty, -1$)	50	19
(-1,0)	30	41
(0,1)	20	41
(1, ∞)	20	19
Total	120	120

We have $\sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \sim \chi_p^2$, where $p = 4 - 1 = 3$. Note we did not estimate parameters, so we don't subtract anything further.

Thank you!

Thank you for this term :) It has been a rewarding and fun experience teaching everyone here !

Please give me feedback in the MyExperience - both positive and negative feedback is welcomed :) It helps me out a lot.

Tutorial Questions

2.5.7, 2020 Final Q5, 2.5.9