All of Statistics - Chapter 10 Solutions

Aug 21, 2020

1.

Let

$$Z_n = \frac{\hat{\theta} - \theta_\star}{\widehat{\text{se}}}.$$

The probability of correctly rejecting the null hypothesis is

$$\begin{aligned}
\beta(\theta_\star) &= 1 - \mathbb{P}(|W| \leq z_{\alpha/2}) \\
&= 1 - \mathbb{P}\left(\frac{\theta_0 - \theta_\star}{\widehat{\text{se}}} - z_{\alpha/2} \leq Z_n \leq \frac{\theta_0 - \theta_\star}{\widehat{\text{se}}} + z_{\alpha/2}\right) \\
&= 1 - \mathbb{P}\left(Z_n \leq \frac{\theta_0 - \theta_\star}{\widehat{\text{se}}} + z_{\alpha/2}\right) + \mathbb{P}\left(Z_n \leq \frac{\theta_0 - \theta_\star}{\widehat{\text{se}}} - z_{\alpha/2}\right)
\end{aligned}$$

If Z_n is asymptotically normal, taking limits in the above yields expression (10.6).

2.

Suppose the conditions of Theorem 10.12 hold and that the CDF F of $T \circ X^n$ is strictly increasing. Then,

$$\mathbb{P}_{\theta_0}[T(X^n) \geq T(x^n)] = 1 - F[T(x^n)]$$

and hence

$$\begin{aligned}
\mathbb{P}_{\theta_0}[\text{p-value} \leq y] &= \mathbb{P}_{\theta_0}[F(T(x^n)) \geq 1 - y] \\
&= 1 - \mathbb{P}_{\theta_0}[T(x^n) \leq F^{-1}(1 - y)] = 1 - F(F^{-1}(1 - y)) = y.
\end{aligned}$$

3.

Recall that the Wald test rejects if and only if $|W| > z_{\alpha/2}$. Equivalently, it does not reject if and only if

$$\hat{\theta} - z_{\alpha/2} \cdot \widehat{\text{se}} \leq \theta_0 \leq \hat{\theta} + z_{\alpha/2} \cdot \widehat{\text{se}}.$$

4.

Note that

$$\text{p-value} = \inf \left\{ \sup_{\theta \in \Theta_0} P_\theta(T(X^n) \geq c_\alpha) : T(x^n) \geq c_\alpha \right\}.$$

Assuming that to each observed test statistic $T(x^n)$ there exists a test with $c_\alpha = T(x^n)$, the infimum above is attained at $c_\alpha = T(x^n)$ and the desired result follows.

5.

a)

The power function is

$$\beta(\theta) = P_\theta(Y > c) = 1 - (c/\theta)^n.$$

b)

See Part (d).

c)

We should not reject the null if we observe 0.48 since

$$\text{p-value} = P_{1/2}(Y \geq 0.48) = 1 - (2 \cdot 0.48)^{20} \approx 0.558.$$

d)

A test of size α is obtained by setting

$$c_\alpha \equiv \frac{1}{2}(1 - \alpha)^{1/n}.$$

converges to monotonically to 0.5 as α converges monotonically to zero from above. Therefore, all possible tests reject the observation 0.52 (since it is greater than 0.5) and hence the corresponding p-value is exactly zero. In this case, we can reject the null with zero probability of making a type I error.

6.

Let $\hat{\theta}$ be the fraction of deaths that occur after passover. Note that either Wald statistic

$$W_0 = \frac{\hat{\theta} - \theta_0}{\sqrt{V_{\theta_0}(\hat{\theta})}} = \sqrt{n}\frac{\hat{\theta} - \theta_0}{\sqrt{\theta_0\left(1 - \theta_0\right)}}$$

or

$$W = \frac{\hat{\theta} - \theta_0}{\widehat{\mathrm{se}}(\hat{\theta})} = \sqrt{n}\frac{\hat{\theta} - \theta_0}{\sqrt{\hat{\theta}\left(1 - \hat{\theta}\right)}}$$

are asymptotically normal under the null hypothesis and hence appropriate (see Remark 10.5). The p-value for the latter is

$$2\Phi(-|w|) = 2\Phi\left(-\left|\sqrt{1919}\frac{997/1919 - 1/2}{(997/1919)(922/1919)}\right|\right) \approx 0.087.$$

This is weak evidence against the null. A 95% confidence interval for the probability of death before passover is

$$922/1919 \pm 2 \cdot \widehat{\mathrm{se}} \approx (0.46, 0.50).$$

7.

a)

Evaluating the code below reveals a p-value of approximately 0.00008 and a 95% confidence interval of approximately (0.01, 0.03).

```
import numpy as np
import scipy.stats

twain     = [.225, .262, .217, .240, .230, .229, .235, .217]
snodgrass = [.209, .205, .196, .210, .202, .207, .224, .223, .220, .201]
delta_hat = np.mean(twain) - np.mean(snodgrass)
var_hat   = np.var(twain) / len(twain) + np.var(snodgrass) / len(snodgrass)
se_hat    = np.sqrt(var_hat)
wald      = delta_hat / se_hat
p_value   = 2. * scipy.stats.norm.cdf(-np.abs(wald))
ci_95_lo  = delta_hat - 2. * se_hat
ci_95_hi  = delta_hat + 2. * se_hat
```

b)

The calculations in Part (a) relied on large sample methods despite there being only a handful of samples. A better choice is a permutation test, which does not require many samples. Such a test is used below to obtain a p-value of approximately 0.0007. This is still very strong evidence against the null.

```
import numpy as np

n_sims = 10**5

def test_stat(data_):
    twain_, snodgrass_ = np.split(data_, [twain.size])
```

```
    return np.abs(np.mean(twain_) - np.mean(snodgrass_))

# Compute test statistic on original data.
twain     = [.225, .262, .217, .240, .230, .229, .235, .217]
snodgrass = [.209, .205, .196, .210, .202, .207, .224, .223, .220, .201]
data      = np.concatenate([twain, snodgrass])
observed  = test_stat(data)

# Repeatedly shuffle and compute test statistic.
np.random.seed(1)
perm_stats = np.empty([n_sims])
for i in range(n_sims):
    np.random.shuffle(data)
    perm_stats[i] = test_stat(data)

p_value = np.sum(perm_stats > observed) / n_sims
```

8.

a)

Let Z be a standard normal random variable. Then, under the null hypothesis,

$$\mathrm{P}_0\left(\frac{X_1 + \cdots + X_n}{n} > c\right) = \mathrm{P}\left(Z > c\sqrt{n}\right) = \Phi(-c\sqrt{n}).$$

Therefore, a test of size α is obtained by taking

$$c = -\frac{\Phi^{-1}(\alpha)}{\sqrt{n}}.$$

b)

If the null hypothesis is false, the power is

$$\beta(1) = \mathrm{P}_1\left(\frac{X_1 + \cdots + X_n}{n} > c\right) = \mathrm{P}\left(Z > (c-1)\sqrt{n}\right) = \Phi(-(c-1)\sqrt{n}).$$

c)

For a fixed size α,

$$\beta(1) = \Phi(\sqrt{n} + \Phi^{-1}(\alpha)).$$

Taking limits yields the desired result.

9.

Let

$$x_n = \frac{\theta_0 - \theta_1}{\hat{\text{se}}}.$$

Then,

$$\beta(\theta_1) = \text{P}_{\theta_1}(|Z| > z_{\alpha/2}) = 1 - \text{P}_{\theta_1}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2})$$
$$= 1 - \text{P}_{\theta_1}(x_n - z_{\alpha/2} \leq Z + x_n \leq x_n + z_{\alpha/2}) = 1 - \Phi(x_n + z_{\alpha/2}) + \Phi(x_n - z_{\alpha/2}).$$

Since $x_n \rightarrow \text{sign}(\theta_0 - \theta_1)\infty$, it follows that $\beta(\theta_1)$ converges to one in both the $\theta_1 > \theta_0$ and $\theta_1 < \theta_0$ case. In other words, as the number of samples increase, the probability of rejection if the null hypothesis is false approaches one.

10.

For each of the four weeks, a separate test is performed. Each test is a paired comparison (Example 10.7) whose null hypothesis is that the rate of death among the two populations is equal. Evaluating the code below yields

Week	**p-value**	**Boneferroni corrected p-value**
-2	0.48	1
-1	0.0046	0.018
1	0.0068	0.027
2	0.27	1

Subject to a Bonferroni correction, there is strong evidence (p-value less than 0.05) to reject the null for weeks -1 and 1.

```
import numpy as np
import scipy.stats

data = np.array([[55, 141],
                 [33, 145],
                 [70, 139],
                 [49, 161]])

totals = np.sum(data, axis=0)
fracs = data / totals
deltas = fracs @ [1., -1.]
std_errs = np.sqrt(np.sum(fracs * (1. - fracs) / totals, axis=1))
wald_stats = deltas / std_errs
p_values = 2. * scipy.stats.norm.cdf(-np.abs(wald_stats))
bonferroni_p_values = np.minimum(p_values.size * p_values, 1.)
```

11.

a)

Drug	**p-value**	**Odds ratio**	**Bonferroni p-value**
Chlorpromazine	0.0057	0.41	0.023
Dimenhydrinate	0.52	1.2	1
Pentobarbital (100 mg)	0.63	0.85	1

Drug	p-value	Odds ratio	Bonferroni p-value
Pentobarbital (150 mg)	0.01	0.56	0.4

The table above is generated by the code below.

```
import numpy as np
import scipy.stats

n_patients = np.array([80, 75, 85, 67, 85])
n_nausea = np.array([45, 26, 52, 35, 37])

fracs = n_nausea / n_patients
variances = fracs * (1. - fracs) / n_patients

odds_ratios = fracs[1:] / fracs[0]
deltas = fracs[1:] - fracs[0]
std_errs = np.sqrt(variances[1:] + variances[0])
wald_stats = deltas / std_errs
p_values = 2. * scipy.stats.norm.cdf(-np.abs(wald_stats))

bonferroni_p_values = np.minimum(p_values.size * p_values, 1.)
```

b)

The Bonferroni p-values are given above.

TODO(BH procedure)

12.

a)

Let $\hat{\lambda} = n^{-1} \sum_n X_n$ be the MLE. Then, $\mathrm{V}(\hat{\lambda}) = n^{-1}\lambda$ and hence $\mathrm{se}(\hat{\lambda}) = \sqrt{n^{-1}\lambda}$. Therefore, a valid Wald statistic is

$$W = \sqrt{n} \frac{\hat{\lambda} - \lambda_0}{\sqrt{\lambda_0}}.$$

The rejection criteria is $|W| > c$. Taking $c = z_{\alpha/2}$ yields a test that has asymptotic size α. Such a rejection region is appropriate when n is large.

For small n, note that

$$\beta(\lambda_0) = \mathrm{P}_{\lambda_0}(|W| > c) = 1 - \mathrm{P}_{\lambda_0}\left(\left|\hat{\lambda} - \lambda_0\right| \leq c\sqrt{\lambda_0 / n}\right)$$
$$= 1 - \mathrm{P}_{\lambda_0}(n\lambda_0 - c\sqrt{n\lambda_0} \leq X_1 + \cdots + X_n \leq n\lambda_0 + c\sqrt{n\lambda_0}).$$

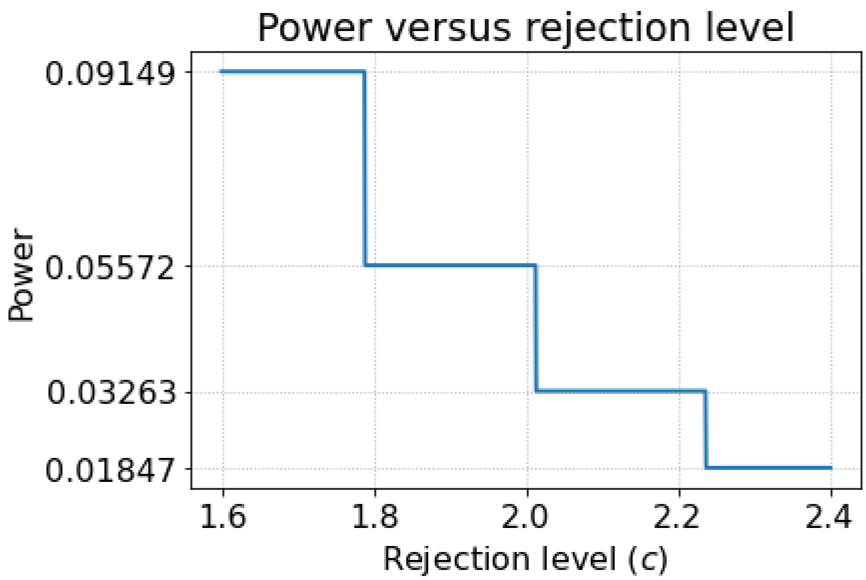
Let $Y = \sum_i X_i \sim \text{Poisson}(n\lambda_0)$. Then,

$$\beta(\lambda_0) = 1 - F_Y((n\lambda_0 + c\sqrt{n\lambda_0}) -) + F_Y(n\lambda_0 - c\sqrt{n\lambda_0}).$$

Finding c such that this quantity is as close to α yields the desired test.

b)

As discussed in Part (a), c is chosen so that the resulting test has power closest to 0.05. This yields a test of power approximately 0.05572. Evaluating the code below, a type I error rate of 0.05578 is observed. If n were larger, a Wald test whose power is closer to 0.05 could be constructed.




```
import numpy as np
import scipy.stats

lambda_0 = 1.
n = 20
alpha = 0.05
n_sims = 10**7
c = scipy.stats.norm.ppf(0.975)  # Approximately 1.96.

np.random.seed(1)
samples = np.random.poisson(lam=lambda_0, size=[n_sims, n])
wald = np.sqrt(n / lambda_0) * (np.mean(samples, axis=1) - lambda_0)
n_reject = np.sum(np.abs(wald) > c)
type_one_err_rate = n_reject / n_sims
```

13.

Recall that

$$\log \mathcal{L} = -\frac{n}{2}\log(2\pi) - \log\sigma - \frac{1}{2\sigma^2}\sum_i \left(X_i - \mu\right)^2.$$

Let $\hat{\mu} = n^{-1}\sum_i X_i$ be the MLE. The likelihood ratio statistic is

$$\lambda = 2\log\mathcal{L}(\hat{\mu}) - 2\log\mathcal{L}(\mu) = \frac{1}{\sigma^2}\left(\sum_i \left(X_i - \mu_0\right)^2 - \left(X_i - \hat{\mu}\right)^2\right)$$

$$= \frac{1}{\sigma^2}\left(n\left(\mu_0^2 - \hat{\mu}^2\right) - 2\left(\mu_0 - \hat{\mu}\right)\sum_i X_i\right) = \frac{n}{\sigma^2}\left(\mu_0^2 + \hat{\mu}^2 - 2\mu_0\hat{\mu}\right) = n\frac{\left(\hat{\mu} - \mu_0\right)^2}{\sigma^2}.$$

The Wald test statistic is

$$W = \frac{\hat{\mu} - \mu_0}{\text{se}(\hat{\mu})} = \sqrt{n}\frac{\left(\hat{\mu} - \mu_0\right)}{\sigma}.$$

Note, in particular, that $W^2 = \lambda$.

14.

The likelihood ratio statistic is

$$\lambda = 2\log\mathcal{L}(\hat{\sigma}) - 2\log\mathcal{L}(\sigma_0) = 2n\left(\log\sigma_0 - \log\hat{\sigma}\right) + \left(\frac{1}{\sigma_0^2} - \frac{1}{\hat{\sigma}^2}\right)\sum_i \left(X_i - \mu\right)^2$$

$$= 2n\left(\log\sigma_0 - \log\hat{\sigma}\right) + n\frac{\hat{\sigma}^2 - \sigma_0^2}{\sigma_0^2}.$$

The Wald test statistic is

$$W = \frac{\hat{\sigma} - \sigma_0}{\widehat{\text{se}}(\hat{\sigma})} = \sqrt{n}\frac{\hat{\sigma} - \sigma_0}{\sqrt{1/I(\hat{\sigma})}} = \sqrt{2n}\frac{\hat{\sigma} - \sigma_0}{\hat{\sigma}}.$$

It is shown in Question 16 that $W^2/\lambda \xrightarrow{P} 1$ under the null hypothesis.

15.

The log likelihood is

$$\log\mathcal{L}(p) = \log\binom{n}{X} + X\log p + (n - X)\log(1 - p).$$

Therefore, the likelihood ratio statistic is

$$\lambda = 2X\left(\log\hat{p} - \log p_0\right) + 2(n - X)\left(\log(1 - \hat{p}) - \log(1 - p_0)\right).$$

The Wald test statistic is

$$W = \sqrt{n}\frac{\hat{p} - p_0}{\sqrt{\hat{p}\left(1 - \hat{p}\right)}}.$$

It is shown in Question 16 that $W^2/\lambda \xrightarrow{P} 1$ under the null hypothesis.

16.

Throughout this proof, it is assumed that the density $f(x; \theta)$ appearing in the likelihood is sufficiently regular. A Taylor expansion reveals

$$\ell(\theta_0) = \ell(\hat{\theta}) + (\hat{\theta} - \theta_0)\ell'(\hat{\theta}) + \frac{1}{2}(\hat{\theta} - \theta_0)^2\ell''(\hat{\theta}) + O((\hat{\theta} - \theta_0)^3).$$

Note, in particular, that $\ell'(\hat{\theta}) = 0$ since $\hat{\theta}$ is an MLE. Therefore,

$$\lambda = 2\log\left(\frac{\mathcal{L}(\hat{\theta})}{\mathcal{L}(\theta_0)}\right) = -(\hat{\theta} - \theta_0)^2\ell''(\hat{\theta}) + O((\hat{\theta} - \theta_0)^3).$$

Moreover,

$$W^2 = \frac{(\hat{\theta} - \theta_0)^2}{\widehat{\mathrm{se}}(\hat{\theta})^2} = nI(\hat{\theta})(\hat{\theta} - \theta_0)^2.$$

It follows that

$$\frac{\lambda}{W^2} = \frac{n^{-1}\ell''(\hat{\theta})}{-I(\hat{\theta})} + O(\hat{\theta} - \theta_0).$$

Under the null hypothesis, $\hat{\theta} \xrightarrow{P} \theta_0$. Therefore, by two applications of Theorem 5.5 (f), $1/I(\hat{\theta}) \to 1/I(\theta_0)$ where

$$I(\theta_0) = \mathrm{E}_{\theta_0}\left[\frac{\partial^2 \log f(X; \theta_0)}{\partial\theta^2}\right].$$

Since

$$\ell''(\theta) = \sum_n \frac{\partial^2 \log f(X_n; \theta)}{\partial\theta^2},$$

by the weak law of large numbers, $n^{-1}\ell''(\hat{\theta}) \xrightarrow{P} I(\theta_0)$ under the null hypothesis. The result now follows by Theorem 5.5 (d).