STA200C HW3

1. Casella 10.1

   First calculate some moments for this distribution.

   $$\mathrm{E}X = \theta/3, \quad \mathrm{E}X^2 = 1/3, \quad \mathrm{Var}X = \frac{1}{3} - \frac{\theta^2}{9}$$

   So $3\bar{X}_n$ is an unbiased estimator of θ with variance

   $$\mathrm{Var}(3\bar{X}_n) = 9(\mathrm{Var}X)/n = (3 - \theta^2)/n \to 0 \text{ as } n \to$$

   So by Casella Theorem 10.1.3, $3\bar{X}_n$ is a consistent estimator of θ.

2. Casella 10.3

   (a) The log likelihood is

   $$-\frac{n}{2}\log(2\pi\theta) - \frac{1}{2}\sum(x_i - \theta)/\theta$$

   Differentiate and set equal to zero, and a little algebra will show that the MLE is the root of $\theta^2 + \theta - W = 0$. The roots of this equation are $(1 - \pm\sqrt{1+4W})/2$, and the MLE is the root with the plus sign, as it has to be nonnegative.

   (b) The second derivative of the log likelihood is $(-2\sum x_i^2 + n\theta)/(2\theta^3)$, yielding an expected Fisher information of

   $$I(\theta) = -\mathrm{E}_\theta \frac{-2\sum X_i^2 + n\theta}{2\theta^3} = \frac{2n\theta + n}{2\theta^2}$$

   and by Casella Theorem 10.1.12 the variance of the MLE is $1/I(\theta)$.

3. Casella 10.4

   (a) Write

   $$\frac{\sum X_i Y_i}{\sum X_i^2} = \frac{\sum X_i(X_i + \epsilon_i)}{\sum X_i^2} = 1 + \frac{\sum X_i \epsilon_i}{\sum X_i^2}.$$

   From normality and independence

   $$EX_i\epsilon_i = 0, \quad \mathrm{Var}X_i\epsilon_i = \sigma^2(\mu^2 + \tau^2), \quad \mathrm{E}X_i^2 = \mu^2 + \tau^2, \quad \mathrm{Var}X_i^2 = 2\tau^2(2\mu^2 + \tau^2),$$

   and $\mathrm{Var}(X_i, X_i\epsilon_i)$. Applying the formulas of Casella Example 5.5.27, the asymptotic mean and variance are

   $$\mathrm{E}\left(\frac{\sum X_i Y_i}{\sum X_i^2}\right) \approx 1 \text{ and } \mathrm{Var}\left(\frac{\sum X_i Y_i}{\sum X_i^2}\right) \approx \frac{n\sigma^2(\mu^2 + \tau^2)}{[n(\mu^2 + \tau^2)]^2} = \frac{\sigma^2}{n(\mu^2 + \tau^2)}$$

(b)

$$\frac{\sum Y_i}{\sum X_i} = \beta + \frac{\sum \epsilon_i}{\sum X_i}$$

with approximate mean β and variance $\sigma^2/(n\mu^2)$.

(c)

$$\frac{1}{n}\sum \frac{Y_i}{X_i} = \beta + \frac{1}{n}\sum \frac{\epsilon_i}{X_i}$$

with approximate mean β and variance $\sigma^2/(n\mu^2)$.

4. Casella 10.5 Errata: " approaches 0" should be " approaches 0 as $n \to \infty$".

   (a) The integral of $\mathrm{E}T_n^2$ is unbounded near zero. We have

$$\mathrm{E}T^2 > \sqrt{\frac{n}{2\pi\sigma^2}} \int_0^1 \frac{1}{x^2} e^{-(x-\mu)^2/2\sigma^2} dx > \sqrt{\frac{n}{2\pi\sigma^2}} K \int_0^1 \frac{1}{x^2} dx = \infty$$

   where $K = \max_{0 \le x \le 1} e^{-(x-\mu)^2/2\sigma^2}$.

   (b) If we delete the interval $(-\delta, \delta)$, then the integrand is bounded, that is, over the range of integration $1/x^2 < 1/\delta^2$.

   (c) Assume $\mu > 0$. A similar argument works for $\mu < 0$. Then

$$P(-\delta < X < \delta) = P[\sqrt{n}(-\delta-\mu) < \sqrt{n}(X-\mu) < \sqrt{n}(\delta-\mu)] < P[Z < \sqrt{n}(\delta-\mu)]$$

   where $Z \sim N(0,1)$. For $\delta < \mu$, the probability goes to 0 as $n \to \infty$.

5. Casella 10.10

   (a) Let $\hat{p} := \sum_{i=1}^n X_i/n$. According to Example 10.1.14 in Casella 2001,

$$\sqrt{n}(\hat{p} - p) \to n(0, p(1-p)) \quad \text{in distribution.}$$

   Let $g(p) = p(1-p)$ and thus $g'(p) = 1 - 2p$ exists and is not 0 when $p \neq \frac{1}{2}$. We then apply the Theorem 5.5.24 (Delta Method) in Casella 2001 and get

$$\sqrt{n}[g(X_n) - g(p)] \to n(0, p(1-p)[1-2p]^2).$$

   And from Example 10.1.14 we know that

$$E_p\left(\left(\frac{\partial}{\partial p}\log f(X|p)\right)^2\right) = \frac{1}{p(1-p)}$$

   Thus we actually have

$$p(1-p)[1-2p]^2 = \frac{[g'(p)]^2}{E_p\left(\left(\frac{\partial}{\partial p}\log f(X|p)\right)^2\right)}.$$

(b) Similarly, applying Theorem 5.5.26 (Second-order Delta Method), $g''(p) = -2$,

$$n[g(Y_n) - g(p)] \to p(1-p)\frac{-2}{2}\chi_1^2 \quad \text{in distribution.}$$

(c) We need to compute the higher moment of binomial variable, since $\hat{p} = Y/n$ where Y is binomial distribution. The table of higher moment can be found in `https://mathworld.wolfram.com/BinomialDistribution.html`. Or you can search "moments of binomial distribution" in Wolfram Alpha `https://www.wolframalpha.com/input/?i=moments+of+binomial+distribution`.

$$\begin{aligned}
\mathrm{E}(\hat{p}(1-\hat{p})) &= \frac{n-1}{n}p(1-p) \\
&= p(1-p) - \frac{p(1-p)}{n} \\
\mathrm{E}(\hat{p}^2(1-\hat{p})^2) &= (p-1)^2p^2 \\
&\quad + \frac{p(1-p)(6p^2-6p+1)}{n^3} \\
&\quad - \frac{p(1-p)(11p^2-11p+2)}{n^2} \\
&\quad + \frac{p(1-p)(6p^2-6p+1)}{n} \\
\mathrm{Var}(\hat{p}(1-\hat{p})) &= \mathrm{E}(\hat{p}^2(1-\hat{p})^2) - (\mathrm{E}(\hat{p}(1-\hat{p})))^2 \\
&= (1-p)p\left(\frac{1-6p+6p^2}{n^3} - 2\frac{1-5p+5p^2}{n^2} + \frac{(1-2p)^2}{n}\right)
\end{aligned}$$

If $p \neq \frac{1}{2}$, it is $O(1/n)$. If $p = \frac{1}{2}$, it is $O(1/n^2)$.
Yes, the reason of the failure of the approximation is clearer, if we have this expression.

6. Casella 10.25

By transforming $y = x - \theta$,

$$\int_{-\infty}^{\infty} \psi(x-\theta)f(x-\theta)dx = \int_{-\infty}^{\infty} \psi(y)f(y)dydy$$

Since ψ is an odd function, $\psi(y) = -\psi(-y)$, and

$$\begin{aligned}
\int_{-\infty}^{\infty} \psi(y)f(y)dy &= \int_{-\infty}^{0} \psi(y)f(y)dy + \int_{0}^{\infty} \psi(y)f(y)dy \\
&= \int_{-\infty}^{0} -\psi(-y)f(y)dy + \int_{0}^{\infty} \psi(y)f(y)dy \\
&= -\int_{-\infty}^{0} \psi(y)f(y)dy + \int_{0}^{\infty} \psi(y)f(y)dy = 0,
\end{aligned}$$

where in the last line we made the transformation $y \to -y$ and used the fact the f is symmetric, so $f(y) = f(-y)$. From the discussion preceding Example 10.2.6 in Casella, $\hat{\theta}_M$ is asymptotically normal with mean equal to the true θ.

7. Casella 10.34

   (a) Let $\hat{p} := \sum_{i=1}^n X_i/n$.

   $-2\log\lambda(\mathbf{x}) := -2(n\hat{p}\log p_0 + n(1-\hat{p})\log(1-p_0) - n\hat{p}\log\hat{p} - n(1-\hat{p})\log(1-\hat{p}))$

   (b) See figure 1 and listing 1.

Listing 1: Simulations

```
sim_engine = function(n, p0) {
    xs = rbinom(n, 1, p0)
    -2*(
        sum(log(dbinom(xs, 1, p0)))
        -sum(log(dbinom(xs, 1, mean(xs))))
    )
}

n = 25 # data size
p0 = 0.3 # p0
N = 10000 # no. of simulation

x_samples = replicate(N, sim_engine(n, p0))
hist(x_samples, freq = F, breaks = 50,
     xlab = "-2log␣lambda", main = "Simulations")
x_d = seq(0, max(x_samples), length = 100)
y_d = dchisq(xfit, df = 1)
lines(xfit, yfit)
```

8. Casella 10.35

   (a) Since $\sigma/\sqrt{n}$ is the estimated standard deviation of $\bar{X}$ in this case, the statistic is a Wald statistic.

   (b) The MLE of σ^2 is $\hat{\sigma}^2_\mu = \sum_i (x_i - \mu)^2/n$ The information number is

$$-\frac{d^2}{d(\sigma^2)^2}\left(-\frac{n}{2}\log\sigma^2 - -\frac{1}{2}\frac{\hat{\sigma^2_\mu}}{\sigma^2}\right)\Bigg|_{\sigma^2=\hat{\sigma}^2_\mu} = \frac{n}{2\hat{\sigma}^2_\mu}$$

   Using the Delta method, the variance of $\hat{\sigma}_\mu = \sqrt{\hat{\sigma}^2_\mu}$ is $\hat{\sigma}^2_\mu/8n$ and a Wald statistic is

$$\frac{\hat{\sigma}_\mu - \sigma_0}{\sqrt{\sigma^2_\mu/8u}}.$$

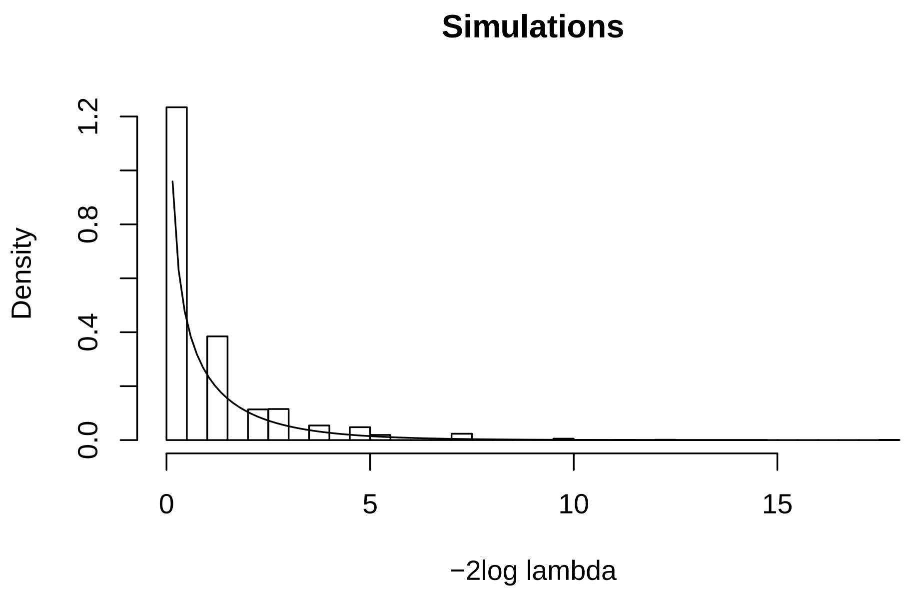

Figure 1: Simulations

9. Casella 10.37

(a) The log likelihood is

$$\log L = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}\sum_i (x_i-\mu)^2/\sigma^2$$

with

$$\frac{d}{d\mu} = \frac{1}{\sigma^2}\sum_i (x_i-\mu) = \frac{n}{\sigma^2}(\bar{x}-\mu)$$
$$\frac{d^2}{d\mu^2} = -\frac{n}{\sigma^2},$$

so the test statistic for the score test is

$$\frac{\frac{n}{\sigma^2}(\bar{x}-\mu)}{\sqrt{\sigma^2/n}} = \sqrt{n}\frac{\bar{x}-\mu}{\sigma}.$$

(b) We test the equivalent hypothesis $H_0 : \sigma^2 = \sigma_0^2$. The likelihood is the same as Casella Exercise 10.35(b), with first derivative

$$-\frac{d}{d\sigma^2} = \frac{n(\hat{\sigma}^2_\mu - \sigma)}{2\sigma^4}$$

and expected information number

$$\mathrm{E}\left(\frac{n(2\hat{\sigma}_\mu^2 - \sigma^2)}{2\sigma^6}\right) = \left(\frac{n(2\sigma^2 - \sigma^2)}{2\sigma^6}\right) = \frac{n}{2\sigma^4}.$$

The score test statistic is

$$\frac{n}{2}\frac{\hat{\sigma}_\mu^2 - \sigma_0^2}{\sigma_0^2}.$$