Math 494: Mathematical Statistics
Solutions to HW2

Problem 1 (# 4.1.3)

(a) The log-likelihood function is given by

$$\ell(\theta) = \log L(\theta) = \log \Pi_{i=1}^n f(x_i;\theta) = \log \Pi_{i=1}^n \frac{e^{-\theta}\theta^{x_i}}{x_i!} = -n\theta + \sum_{i=1}^n x_i \log\theta + \text{Const.}$$

By solving

$$0 = \left.\frac{\partial \ell(\theta)}{\theta}\right|_{\hat{\theta}_{MLE}} = -n + \sum_{i=1}^n \frac{x_i}{\theta},$$

we have

$$\hat{\theta}_{MLE} = \frac{1}{n}\sum_{i=1}^n x_i = \bar{x}.$$

The MLE is unbiased since $\mathbb{E}(\frac{1}{n}\sum_{i=1}^n x_i) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(x_i) = \theta$.

(b) A realization of the estimator given the data is

$$\hat{\theta} = \frac{9+7+9+15+10+13+11+7+2+12}{10} = 9.5.$$

This means based on the observed data, we would expect to see about 9.5 customers coming into the store between 9:00 am to 10:00 am.

Problem 2 (# 4.1.5)

(a) Using conditional expectation we have

$$\begin{aligned} P(X_1 \le X_i, i = 2,3,\ldots,j) &= \mathbb{E}[P(X_1 \le X_i, i=2,3,\ldots,j|X_1)] \\ &= \mathbb{E}[(1-F(X_1))^{j-1}] = \int_0^1 u^{j-1}du = j^{-1}, \end{aligned}$$

where we used the fact that the random variable $F(X_1)$ has a uniform (0,1) distribution.

(b) In the same way, for $j = 2,3,\ldots$

$$\begin{aligned} P(Y = j-1) &= P(X_1 \le X_2, \ldots, X_1 \le X_{j-1} X_j > X_1) \\ &= \mathbb{E}[(1-F(X_1))^{j-2}F(X_1)|X_1] = \int_0^1 u^{j-2}(1-u)du \\ &= \frac{1}{j(j-1)}. \end{aligned}$$

(c)

$$\mathbb{E}(Y=y) = \sum_{y=1}^\infty y\frac{1}{y(y+1)} = \sum_{y=1}^\infty \frac{1}{y+1} \to \infty.$$

So the mean does not exist, and hence the variance does not exist either.

Problem 3 (# 4.2.4)

(a) $X_1, X_2, \cdots, X_n \sim^{iid} \Gamma(1, \theta)$, so for any $i = 1, \cdots, n$, the moment generating function (MGF) for X_i is

$$m(t) = (1 - \theta t)^{-1}.$$

Then the MGF for $(2/\theta)X_i$ is

$$E[e^{(2t/\theta)X_i}] = m(2t/\theta) = (1 - 2t)^{-1},$$

which is the MGF for a χ^2 distribution with 2 degrees of freedom.

Because of additivity of χ^2 distributions, $(2/\theta)\sum_{i=1}^n X_i = \sum_{i=1}^n (2/\theta)X_i$ has a χ^2 distribution with $2n$ degrees of freedom.

(b) Let U be a random variable that has a χ^2-distribution with k degrees of freedom, and $u_{\alpha,k}$ be the real number for which $P(U > u_{\alpha,k}) = \alpha$.

Here, $U = (2/\theta)\sum_{i=1}^n X_i$ and $k = 2n$. Use the fact

$$P(u_{1-\alpha/2,2n} < (2/\theta)\sum_{i=1}^n X_i < u_{\alpha/2,2n})) = 1 - \alpha$$

to derive the two-sided $(1 - \alpha)100\%$ confidence interval for θ

$$\left(\frac{2}{u_{\alpha/2,2n}} \sum_{i=1}^n X_i,\ \frac{2}{u_{1-\alpha/2,2n}} \sum_{i=1}^n X_i \right).$$

(c) In problem 4.2.2, $n = 20$ and $\sum_{i=1}^n X_i = 2023$. $\alpha = 0.05$, and the quantiles

$$u_{0.025,40} = 59.3417, u_{0.975,40} = 24.4330.$$

The 95% confidence interval for θ is $\left(\frac{2}{59.3417}2023, \frac{2}{24.4330}2023\right) = (68.181, 165.595)$, which is longer than the approximate large sample confidence interval $(54.953, 147.347)$ from problem 4.2.2. It is worth noting that $\overline{X}$ is only approximately normal, but actually $(2/\theta)\sum_{i=1}^{20} X_i$ has exactly a χ^2 distribution with 40 degrees of freedom.

Problem 4 (# 4.2.10)

(a) $\sqrt{9}(\overline{X} - \mu)/\sigma \sim N(0, 1)$, and hence the 95% confidence interval for μ is

$$\left(\overline{X} \pm z_{0.025}\sigma/\sqrt{9}\right) = \left(\overline{X} \pm 1.96\sigma/3\right).$$

Therefore, the length of the confidence interval is

$$(2)(1.96\sigma)/3 = 1.3067\sigma.$$

(b) $\sqrt{9}(\overline{X}-\mu)/S \sim t(8)$, and hence the 95% confidence interval for μ is

$$\left(\overline{X} \pm t_{0.025,8} S/\sqrt{9}\right) = \left(\overline{X} \pm 2.3060 S/3\right).$$

Therefore, the expected length of this confidence interval is

$$(2)(2.306 E[S])/3 = 1.5373 E[S].$$

Use the fact $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$, we have $8S^2/\sigma^2 \sim \chi^2(8)$. By Theorem 3.3.1,

$$\frac{\sqrt{8}}{\sigma} E\left[S\right] = \frac{\sqrt{2}\Gamma(4.5)}{\Gamma(4)} = 2.7416.$$

Therefore, $E[S] = 0.9693\sigma$ and the expected length is $1.5373 E[S] = 1.4901\sigma$.

(c) On average, the confidence intervals for μ when σ is unknown are longer than those when σ is known.

Problem 5 (# 4.2.18)

(a) The statement follows

$$a < (n-1)S^2/\sigma^2 < b \Leftrightarrow a\sigma^2 < (n-1)S^2 < b\sigma^2 \Leftrightarrow (n-1)S^2/b < \sigma^2 < (n-1)S^2/a$$

(b) For $n = 8$, $a = \chi^2_{.025} = 17.53, b = \chi^2_{.975} = 2.18$, and then,

$$(n-1)S^2/a = 8 * 7.93/17.53 = 3.62, (n-1)S^2/b = 29.10.$$

Therefore, the confidence interval is $[3.62, 29.10]$.

(c) Use the fact that $\Sigma(X_i - \mu)^2/\sigma^2$ is $\chi^2(n)$.

Math 494: Mathematical Statistics
Solutions to HW3

Problem 1 (# 4.4.6)

(a) Let m be the median of the distribution with pdf $f(x) = 2x$, $0 < x < 1$.

$$\int_0^m 2x dx = 0.5 \Rightarrow \; m^2 = 0.5 \Rightarrow \; m = \frac{\sqrt{2}}{2}.$$

Then the probability that the smallest of X_1, X_2, X_3 exceeds the median is

$$\begin{aligned} P\left(\min(X_1, X_2, X_3) > \sqrt{2}/2\right) &= P(X_1 > \sqrt{2}/2, X_2 > \sqrt{2}/2, X_3 > \sqrt{2}/2) \\ &= P(X_1 > \sqrt{2}/2)P(X_2 > \sqrt{2}/2)P(X_3 > \sqrt{2}/2) \;\; \because \text{independence} \end{aligned}$$

The cdf $F(x) = \int_0^x 2tdt = x^2 \Rightarrow \; P\left(X_i > \sqrt{2}/2\right) = 1 - F(\sqrt{2}/2) = 1/2$, for $i = 1, 2, 3$. Therefore, $P\left(\min(X_1, X_2, X_3) > \sqrt{2}/2\right) = 1/8$.

(b) Apply equation (4.4.2) and (4.4.3) on page 229, then the pdf of Y_2 is

$$g_2(y_2) = \begin{cases} 12(y_2^3 - y_2^5), & \text{if } 0 < y_2 < 1 \\ 0, & \text{otherwise} \end{cases}$$

the pdf of Y_3 is

$$g_3(y_3) = \begin{cases} 6y_3^5, & \text{if } 0 < y_3 < 1 \\ 0, & \text{otherwise} \end{cases}$$

and the joint pdf of Y_2, Y_3 is

$$g_{23}(y_2, y_3) = \begin{cases} 24y_2^3 y_3, & \text{if } 0 < y_2 < y_3 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$E[Y_2] = \int_0^1 y_2 g_2(y_2) dy_2 = \int_0^1 t \cdot 12(t^3 - t^5)dt = \frac{12}{5} - \frac{12}{7} = \frac{24}{35}$$

$$E[Y_3] = \int_0^1 y_3 g_3(y_3) dy_2 = \int_0^1 t \cdot 6t^5 dt = \frac{6}{7}$$

$$E[Y_2^2] = \int_0^1 y_2^2 g_2(y_2) dy_2 = \int_0^1 t^2 \cdot 12(t^3 - t^5)dt = 2 - \frac{12}{8} = \frac{1}{2}$$

$$E[Y_3^2] = \int_0^1 y_3^2 g_3(y_3) dy_2 = \int_0^1 t^2 \cdot 6t^5 dt = \frac{3}{4}$$

$$E[Y_2 Y_3] = \int_0^1 \int_0^{y_3} y_2 y_3 g_{23}(y_2, y_3) dy_2 dy_3 = 24 \int_0^1 y_3^2 \left(\int_0^{y_3} y_2^4 dy_2\right) dy_3 = \frac{3}{5}$$

Then

$$Var[Y_2] = E[Y_2^2] - (E[Y_2])^2 = \frac{1}{2} - \left(\frac{24}{35}\right)^2$$

$$Var[Y_3] = E[Y_3^2] - (E[Y_3])^2 = \frac{3}{4} - \left(\frac{6}{7}\right)^2$$

$$Cov(Y_2, Y_3) = E[Y_2 Y_3] - E[Y_2]E[Y_3] = \frac{3}{5} - \frac{24}{35} \cdot \frac{6}{7}$$

Therefore,

$$Corr(Y_2, Y_3) = \frac{Cov(Y_2, Y_3)}{\sqrt{Var[Y_2]} \cdot \sqrt{Var[Y_3]}} = 0.57$$

Problem 2 (# 4.4.8)

The pdf $f(x) = e^{-x}$, $x > 0 \Rightarrow$ The cdf $F(x) = 1 - e^{-x}$, $x > 0$.
The joint pdf of Y_2 and Y_4 is

$$g(y_2, y_4) = \begin{cases} 120[1 - e^{-y_2}][e^{-y_2} - e^{-y_4}]e^{-y_2}e^{-2y_4}, & \text{if } y_4 > y_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

$Z_1 = Y_2$, $Z_2 = Y_4 - Y_2 \ \Rightarrow \ Y_2 = Z_1$, $Y_4 = Z_1 + Z_2$, so $J = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$.

The joint pdf of Z_1 and Z_2 is

$$\begin{aligned} h(z_1, z_2) &= g(z_1, z_1 + z_2)|J| \\ &= \begin{cases} 120[1 - e^{-z_1}][e^{-z_1} - e^{-(z_1+z_2)}]e^{-z_1}e^{-2(z_1+z_2)}, & \text{if } z_1 + z_2 > z_1 > 0 \\ 0, & \text{otherwise} \end{cases} \\ &= \begin{cases} 120[(1 - e^{-z_1})e^{-4z_1}][(1 - e^{-z_2})e^{-2z_2}], & \text{if } z_1 > 0, z_2 > 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

The density is separable, so Z_1 and Z_2 are independent.

Problem 3 (# 4.4.17)

(a) The pdf $f(x) = 2x$, $0 < x < 1$. $\Rightarrow$ The cdf $F(x) = x^2$, $0 < x < 1$.
The joint pdf of Y_3 and Y_4 is

$$g(y_3, y_4) = \begin{cases} 12[F(y_3)]^2 f(y_3) f(y_4) = 48{y_3}^5 y_4, & \text{if } 0 < y_3 < y_4 < 1 \\ 0, & \text{otherwise} \end{cases}$$

(b) First, the marginal pdf of Y_4 is

$$g(y_4) = \begin{cases} 4[F(y_4)]^3 f(y_4) = 8{y_4}^7, & \text{if } 0 < y_4 < 1 \\ 0, & \text{otherwise} \end{cases}$$

Then, the conditional pdf of Y_3, given $Y_4 = y_4$ can be computed by

$$g(y_3|Y_4 = y_4) = \frac{g(y_3, y_4)}{g(y_4)} = \begin{cases} \frac{48{y_3}^5 y_4}{8{y_4}^7} = 6{y_3}^5{y_4}^{-6}, & \text{if } 0 < y_3 < y_4 < 1 \\ 0, & \text{otherwise} \end{cases}$$

(c) The evaluation follows

$$E(Y_3|y_4) = \int_0^{y_4} g(y_3|Y_4 = y_4) y_3 dy_3 = \int_0^{y_4} 6{y_3}^6{y_4}^{-6} dy_3 = \frac{6}{7} y_4$$

Problem 4 (# 4.4.23)

(a) Set $U = X_1 + X_2$, $V = X_2 \Rightarrow X_1 = U - V$, $X_2 = V$, so $J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$
The joint pdf of U and V is

$$f(u, v) = f(x_1, x_2)|J| = f(x_1)f(x_2) = f(u - x_2)f(x_2)$$

The second equation holds because X_1, X_2 are independent.
Then, the pdf of U is

$$f(u) = \int f(u, v)dv = \int f(u - x_2)f(x_2)dx_2$$

Similarly, set $Y = U + X_3$, $W = X_3$, we get the joint pdf of Y and W is

$$f(y, w) = f(y - x_3)f(x_3)$$

And the pdf of Y is

$$f(y) = \int f(y, w)dw = \int f(y - x_3)f(x_3)dx_3 = \int\int f(y - x_2 - x_3)f(x_2)f(x_3)dx_2dx_3$$

Since $f(x_1) = f(x_2) = f(x_3) = 1$ with $0 < x_1, x_2, x_3 < 1$, next, we need to consider the limit of integration for $f(y)$.
For $0 < y \leq 1$,

$$f(y) = \int_0^y \int_0^{y-x_2} 1 dx_3 dx_2 = \int_0^y (y - x_2)dx_2 = \frac{1}{2}y^2$$

For $1 < y \leq 2$,

$$f(y) = \int_0^{y-1} \int_{y-1-x_2}^1 1 dx_3 dx_2 + \int_{y-1}^1 \int_0^{y-x_2} 1 dx_3 dx_2 = -y^2 + 3y - \frac{3}{2}$$

For $2 < y < 3$,

$$f(y) = \int_{y-2}^1 \int_{y-1-x_2}^1 1 dx_3 dx_2 = \int_{y-2}^1 (2 - y + x_2)dx_2 = \frac{y^2}{2} - 3y + \frac{9}{2}$$

This completes the final result.

(b) Sort X_1, X_2, X_3 from small to large, and relable them as Y_1, Y_2, Y_3. Then $Z = Y_3$, and apply equation (4.4.2), we have the pdf of Z is

$$f(z) = f(y_3) = \frac{3!}{2!}[F(y_3)]^2 f(y_3) = 3{y_3}^2 = 3z^2,$$

for $0 < z < 1$, and 0 elsewhere.

Problem 5 (# 4.4.31)

(a) The c.d.f. of the distribution is

$$F(x) = \int_0^\theta \frac{3x^2}{\theta^3} dx = \frac{x^3}{\theta^3}.$$

Then the p.d.f. of Y_n is

$$g_n(y_n) = \frac{n!}{(n-1)!} F^{n-1}(y_n) f(y_n) = \frac{3y_n^{3n-1} n}{\theta^{3n}}$$

given by (4.4.2) on textbook. So

$$P(c < Y_n/\theta < 1) = P(c\theta < Y_n < \theta) = \int_{c\theta}^{\theta} \frac{3y_n^{3n-1} n}{\theta^{3n}} = 1 - c^{3n}.$$

(b) $P(c\theta < Y_4 < \theta) = 1 - c^{12} = 0.95 \ \Rightarrow \ c = (0.05)^{1/12}$. So $P(Y_4 < \theta < Y_4/(0.05)^{1/12}) = 0.95$, then a 95% CI of θ is given by $(2.3, 2.3/c) = (2.3, 2.952)$.

Math 494: Mathematical Statistics
Solutions to HW4

Problem 1 (# 4.5.4)

X has a binomial distribution with pmf $P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}$.
Consider the hypotheses

$$H_0: p = \frac{1}{2} \text{ versus } H_1: p = \frac{1}{4}$$

The rejection rule is given by

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } X_1 \leq 3,$$

Therefore, the significant level α is

$$\begin{aligned}\alpha &= P_{H_0}[X_1 \leq 3] \\ &= P_{H_0}[X_1=0]+P_{H_0}[X_1=1]+P_{H_0}[X_1=2]+P_{H_0}[X_1=3] \\ &= \binom{10}{0}(\frac{1}{2})^0(\frac{1}{2})^{10-0}+\binom{10}{1}(\frac{1}{2})^1(\frac{1}{2})^{10-1}+\binom{10}{2}(\frac{1}{2})^2(\frac{1}{2})^{10-2}+\binom{10}{3}(\frac{1}{2})^3(\frac{1}{2})^{10-3} \quad = 0.17\end{aligned}$$

The power γ is

$$\begin{aligned}\gamma(p) &= P_{H_1}[X_1 \leq 3] \\ &= P_{H_1}[X_1=0]+P_{H_1}[X_1=1]+P_{H_1}[X_1=2]+P_{H_1}[X_1=3] \\ &= \binom{10}{0}(\frac{1}{4})^0(\frac{3}{4})^{10-0}+\binom{10}{1}(\frac{1}{4})^1(\frac{3}{4})^{10-1}+\binom{10}{2}(\frac{1}{4})^2(\frac{3}{4})^{10-2}+\binom{10}{3}(\frac{1}{4})^3(\frac{3}{4})^{10-3} \quad = 0.78\end{aligned}$$

Problem 2 (# 4.5.12)

$Y=\sum_{i=1}^{8} X_i$ follows a Poisson distribution with mean 8μ.
Consider the hypotheses

$$H_0: \mu = 0.5 \text{ versus } H_1: \mu > 0.5$$

The rejection rule is given by

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } Y \geq 8,$$

(a) The significant level α is

$$\alpha = P_{H_0}[Y \geq 8] = 1 - P_{H_0}[Y < 8] = 0.05$$

(b) The power $\gamma(\mu)$ is

$$\gamma(\mu) = P_{H_1}[Y \geq 8] = 1 - P_{H_1}[Y \leq 7] = 1 - e^{-8\mu}\sum_{k=0}^{7}\frac{(8\mu)^k}{k!}$$

(c) Using Table I from Appendix C, we can find

$$\gamma(0.75) = P(Poisson(6) \geq 8) = 0.26$$

$$\gamma(1) = P(Poisson(8) \geq 8) = 0.55$$

$$\gamma(1.25) = P(Poisson(10) \geq 8) = 0.78$$

Problem 3 (# 4.6.8)

(a) $H_0 \ : \ p = 0.14$; $H_1 \ p > 0.14$;

(b) By CLT, under H_0, $\frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$ is asymptotically standard normal where $\hat{p} = y/n$. Then a critical region at level $\alpha = 0.01$ is given by

$$C = \{z : \ z \geq 2.326\} \text{ where } z = \frac{y/n - p_0}{\sqrt{p_0(1-p_0)/n}}.$$

(c) $z = \frac{104/590-0.14}{\sqrt{(0.14)(0.86)/590}} = 2.539 > 2.326$ (with p-value $= 1 - \Phi(2.539) \approx 0.0056 < 0.01$), so reject H_0 and conclude that the campaign was successful.

Problem 4 (# 4.7.3)

Use Chi-square test. $Q_5 = \frac{(b-20)^2}{20} + \frac{(40-b-20}{20} = \frac{(b-20)^2}{10} = 12.8$, which is the 97.5 percentile of the $\chi^2(5)$ distribution. Thus $(b-20)^2 = 128$ and $b = 20 \pm 11.3$. Hence $b < 8.7$ or $b > 31.3$ would lead to rejection.

Problem 5 (# 4.7.7)

If p is known, then $\sum_{i=1}^{3} \frac{(X_i - np_i)^2}{np_i} \sim \chi^2(2)$. If p is unknown, need to first estimate p. The MLE for p is defined by maximizing the likelihood

$$\frac{n}{x_1!x_2!x_3!}[p^2]^{x_1}[2p(1-p)]^{x_2}[(1-p)^2]^{x_3}. \Longrightarrow \text{MLE } \hat{p} = \frac{2X_1 + X_2}{2(X_1 + X_2 + X_3)}.$$

Then $\hat{p}_1 = \hat{p}^2, \hat{p}_2 = 2\hat{p}(1-\hat{p})$, $\hat{p}_3 = (1-\hat{p})^2$, and $\sum_{i=1}^{3} \frac{(X_i - n\hat{p}_i)^2}{n\hat{p}_i}$ has an approximate $\chi^2(1)$ since we estimate a parameter.

Math 494: Mathematical Statistics
Solutions to HW5

Problem 1 (# 4.7.4)

$$H_0 \; : \; p_1 = \frac{9}{16}, \; p_2 = \frac{3}{16}, \; p_3 = \frac{3}{16}, \; p_4 = \frac{1}{16}; \quad H_1 \; : \; \text{not } H_0.$$

Use chi-square test. Let O_j be the observed frequencies and $E_j = np_i$ be the expected frequencies, then

$$Q_3 = \sum_{j=1}^{3} \frac{(O_j - E_j)^2}{E_j} \sim \chi_3^2.$$

The observed value of Q_3 is

$$\frac{(86 - 160 \times \frac{9}{16})^2}{160 \times \frac{9}{16}} + \frac{(35 - 160 \times \frac{3}{16})^2}{160 \times \frac{3}{16}} + \frac{(26 - 160 \times \frac{3}{16})^2}{160 \times \frac{3}{16}} + \frac{(13 - 160 \times \frac{1}{16})^2}{160 \times \frac{1}{16}} \approx 2.44 < \chi^2_{3,0.99} = 11.345.$$

So at the significance level $\alpha = 0.01$, we do not reject the null hypothesis that the data are consistent with the Mendelian theory.

Problem 2 (# 4.7.6)

$$H_0 \; : \; P(A_i \cap B_j) = P(A_i)P(B_j), \; \forall i, j, \; \text{(or, A and B are independent).}$$

$$H_1 \; : \; \text{not } H_0 \text{ (or, A and B are not independent).}$$

The test statistic is

$$Q = \sum_{i=1}^{3} \sum_{j=1}^{4} \frac{(X_{ij} - np_{i\cdot}p_{\cdot j})^2}{np_{i\cdot}p_{\cdot j}} \sim \chi^2_{(4-1)(3-1)} = \chi^2_6.$$

Plugging in

$$P_{1\cdot} = \frac{10 + 21 + 15 + 6}{200} = 0.26, \; P_{2\cdot} = \frac{11 + 27 + 21 + 13}{200} = 0.36,$$

$$P_{3\cdot} = 0.38, \; P_{\cdot 1} = 0.135, \; P_{\cdot 2} = 0.335, \; P_{\cdot 3} = 0.315, \; P_{\cdot 4} = 0.215,$$

we have the observed statistic

$$Q^* = 12.941 > \chi^2_{0.95} \approx 12.6.$$

So at the significance level $\alpha = 0.05$, we reject the null hypothesis and conclude that the attributes of A and B are not independent.

Problem 3 (# 4.7.9)

R code :

```
> freq_obs=c(20,40,16,18,6)
> n=sum(freq_obs)
> x = rep(0:4, times=freq_obs)
> table(x)
x
 0  1  2  3  4
20 40 16 18  6
> probs = dpois(0:3, lambda=mean(x))
> probs=c(probs,1-sum(probs))
> freq_est=n*probs
> freq_est
[1] 22.313016 33.469524 25.102143 12.551072  6.564245
> Q=sum((freq_obs-freq_est)^2/freq_est)
> Q
[1] 7.228557
> qchisq(0.95,df=3)
[1] 7.814728
```

(a) The chi-square goodness-of-fit statistic $Q = 7.23$.

(b) degree of freedom 5-1-1=3

(c) $Q < \chi^2_{0.95,3} = 7.81$, so we accept the null hypothesis.

Problem 4 (# 5.1.2)

(a) $\forall \varepsilon > 0$, by Chebyshev's inequality,

$$P[|Y_n/n - p| \geq \varepsilon] = P\left[|Y_n/n - E[Y_n/n]| \geq \varepsilon\right] \leq \frac{Var(Y_n/n)}{\varepsilon^2} = \frac{np(1-p)}{n^2\varepsilon^2} \to 0\,, \text{ as } n \to \infty.$$

(b) It follows from Theorem 5.1.2 and 5.1.3.

(c) Use the result from part (b). It follows from Theorem 5.1.5.

Problem 5 (# 5.1.6)

By the Weak Law of Large Numbers, the sample mean,$\overline{X}_n$, is a consistent estimator of μ. i.e.$\overline{X}_n \xrightarrow{P} \mu$. And we can show the following as Example 5.1.1:

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \frac{n}{n-1}(\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \overline{X}_n^2)$$
$$\xrightarrow{P} 1.[E(X_1^2) - \mu^2] = \sigma^2$$

So, $S_1^2 \xrightarrow{P} \sigma_1^2, S_2^2 \xrightarrow{P} \sigma_2^2$. And by Theorem 5.1.2, 5.1.3, we have

$$\frac{S_1^2}{n_1} \xrightarrow{P} \frac{\sigma_1^2}{n_1}, \frac{S_2^2}{n_2} \xrightarrow{P} \frac{\sigma_2^2}{n_2}, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \xrightarrow{P} \frac{\sigma_1^2}{n_1} + \frac{S_2^2}{n_2}$$

Therefore, we find that

$$\frac{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}{\frac{\sigma_1^2}{n_1} + \frac{S_2^2}{n_2}} \xrightarrow{P} 1$$

Since the function $g(x) = \sqrt{x}$ is continuous at $x = 1$, we can apply Theorem 5.1.4, and the convergence in probability follows.

$$\frac{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{S_2^2}{n_2}}} \xrightarrow{P} 1$$

Math 494: Mathematical Statistics
Solutions to HW6

Problem 1 (# 5.2.2)

The p.d.f. of y_1 is given by

$$g_1(y_1) = ne^{-n(y_1-\theta)},\ 0 < y_1 < \infty.$$

Since $z = n(y_1 - \theta)$, $\frac{dy_1}{dz} = \frac{1}{n}$, the p.d.f. of z is

$$h_n(z) = e^{-z}.$$

Then the c.d.f. of z is

$$H_n(z) = 1 - e^{-z} \xrightarrow{n\to\infty} \begin{cases} 1 - e^{-z} & 0 < z < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Problem 2 (# 5.2.17)

By the conclusion of # **5.2.16 (a)**,

$$\lim_{n\to\infty} M_{Y_n}(t) = \lim_{n\to\infty}\left[\left(1 + \frac{t}{\sqrt{n}} + \frac{(t/\sqrt{n})^2}{2!} + \frac{(t/\sqrt{n})^3}{3!} + \dots\right) - \left(\frac{t}{\sqrt{n}} - (t/\sqrt{n})^2 - \frac{(t/\sqrt{n})^3}{2!} + \dots\right)\right]^{-n}$$

$$= \lim_{n\to\infty}\left(1 - \frac{t^2}{2n} - \frac{t^3}{3n^{2/3}} - \dots\right)^{-n} = \lim_{n\to\infty}\left(1 - \frac{t^2}{2n}\right)^{-n} = \lim_{n\to\infty}\left(1 - \frac{t^2}{2n}\right)^{-\frac{2n}{t^2}\cdot\frac{t^2}{2}} = e^{t^2/2},$$

which is the MGF of the standard normal distribution, so Y_n's limiting distribution is $\mathcal{N}(0,1)$. Then by the delta's method with $g(X_n) = \sqrt{X_n}$ and $g'(\theta) = \frac{1}{2\sqrt{\theta}}$ where $\theta = 1$,

$$\sqrt{n}(\sqrt{\bar{X}_n} - 1) \xrightarrow{d} N(0, g'^2(1)\cdot 1) = \mathcal{N}\left(0, \frac{1}{4}\right).$$

Problem 3 (# 5.2.18)

$$\begin{aligned} P[Z_n \le t] &= P[Y_n \le t + \log n] \\ &= (P[X_1 \le t + \log n])^n \\ &= \left(1 - e^{-(t+\log n)}\right)^n \\ &= \left(1 - \frac{1}{n}e^{-t}\right)^n \\ &= \left[\left(1 - \frac{e^{-t}}{n}\right)^{-\frac{n}{e^{-t}}}\right]^{-e^{-t}} \to \exp\left\{-e^{-t}\right\}\ ,\quad \text{as } n \to \infty. \end{aligned}$$

The limiting distribution is the standard Gumbel distribution.

Problem 4 (# 5.3.11)

By Theorem 5.2.9, $u(\overline{X})$ is approximately distributed as $N\left(u(\mu), [u'(\mu)]^2\sigma^2/n\right)$.
Here, $u(\overline{X}) = \overline{X}^3$, so $u(\mu) = \mu^3$, $u'(\mu) = 3\mu^2$.
Therefore, the approximate distribution of $\overline{X}^3$ is $N\left(\mu^3, 9\mu^4\sigma^2/n\right)$.

Problem 5 (# 5.3.12)

Since Y/n converges in probability to μ, we can approximate $u(Y/n)$ by the first two terms of its Taylor's expansion about μ, namely, by

$$u(\frac{Y}{n}) \doteq v(\frac{Y}{n}) = u(\mu) + u'(\mu)(\frac{Y}{n} - \mu)$$

$v(\frac{Y}{n})$ is a linear function of Y/n, and thus has an approximate normal distribution; clearly, it has mean $\mu(p)$ and variance

$$[u'(\mu)]^2\frac{\mu}{n}$$

And it is the latter that we want to be free of μ; thus, we have

$$u'(\mu) = \frac{c_1}{\sqrt{\mu}}$$

A solution of this is $u(\mu) = \sqrt{\mu}$, with $c_1 = 1/2$. Therefore, we have $u(\frac{Y}{n}) = \sqrt{\frac{Y}{n}}$.

Math 494: Mathematical Statistics
Solutions to Midterm1

1. (12 points) Assuming X is a discrete random variable whose pmt is $p(x)$ supported on a finite set $\{a_1, \dots, a_m\}$. Let the quantity of interest be $\theta_j = p(a_j)$, $j = 1, \dots, m$. Let $X_1, \dots, X_n$ be a random sample from the distribution with pmf $p(x)$.

   (a) (6 points) Show that the estimator $\hat{\theta}_j = \frac{1}{n}\sum_{i=1}^n 1(X_i = a_j)$ is an unbiased estimator of θ_j. Here $1(A) = 1$ if A is true and 0 otherwise.

   (b) (6 points) Find the variance of $\hat{\theta}_j$.

Solution:

(a) [1pt] We need to show that $E(\hat{\theta}_j) = \theta_j$.

$$
\begin{aligned}
\text{[1pt]} \quad \mathbb{E}(\hat{\theta}_j) &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\big(1(X_i = a_j)\big) \\
\text{[2pts]} \quad &= \frac{1}{n}\sum_{i=1}^n P(X_i = a_j) \\
\text{[2pts]} \quad &= \frac{1}{n}\sum_{i=1}^n p(a_j) = \frac{1}{n}\sum_{i=1}^n \theta_j = \theta_j.
\end{aligned}
$$

(b) Since $1(X_i = a_j)$'s, $i = 1, 2, \dots, n$, are i.i.d., we have

$$\text{[2pts]} \quad Var(\hat{\theta}_j) = \frac{1}{n^2}\sum_{i=1}^n Var\big(1(X_i = a_j)\big).$$

For each i,

$$\text{[2pts]} \quad Var\big(1(X_i = a_j)\big) = p(a_j)(1 - p(a_j)) = \theta_j(1 - \theta_j)$$

since $1(X_i = a_j)$ is a binomial random variable with $p = p(a_j)$. So

$$\text{[2pts]} \quad Var(\hat{\theta}_j) = \frac{\theta_j(1 - \theta_j)}{n}.$$

2. (10 points) Let $\bar{X}$ be the mean of a random sample of size n from the distribution $N(\mu, 4)$. Find n such that $P(\bar{X} - 1 < \mu < \bar{X} + 1) = 0.95$. Represent your answer by the notation Z_α and give the proper value of α. (Note that $P(Z > Z_\alpha) = \alpha$ for a standard normal random variable Z.)

Solution: [2pts] $\overline{X} \sim N(\mu, \frac{4}{n})$, so

$$[4\text{pts}] \qquad P\left[\left|\frac{\overline{X} - \mu}{2/\sqrt{n}}\right| < z_{0.025}\right] = 0.95.$$

This is $P\left[\overline{X} - \frac{2}{\sqrt{n}} z_{0.025} < \mu < \overline{X} + \frac{2}{\sqrt{n}} z_{0.025}\right] = 0.95$.
[4pts] Let $\frac{2}{\sqrt{n}} z_{0.025} = 1$, we find $n = 4z_{0.025}^2$.

3. (20 points) Let $Y_1 < Y_2$ be the order statistics of a random sample of size 2 from $U(0, 1)$.
   (a) (10 points) Find $E(Y_1)$.
   (b) (10 points) Find the covariance of Y_1 and Y_2.

Solution:

(a) Using (4.4.2) in textbook, the p.d.f. of Y_1 is

$$[5\text{pts}] \quad g(y_1) = \frac{2!}{0!1!}[F(y_1)]^0(1 - F(y_1))^1 f(y_1) = 2(1 - y_1),$$

where $F(\cdot)$ and $f(\cdot)$ are the c.d.f. and p.d.f. of the $U(0, 1)$ distribution. So

$$[5\text{pts}] \quad \mathbb{E}(Y_1) = \int_0^1 2(1 - y_1)y_1 dy_1 = \frac{1}{3}.$$

(b) Using (4.4.3) in textbook, the joint p.d.f. of Y_1 and Y_2 is

$$[2\text{pts}] \quad g_{12}(Y_1, Y_2) = \frac{2!}{0!0!0!}[F(y_1)]^0[F(y_2) - F(y_1)]^0[1 - F(y_2)]^0 f(y_1)f(y_2) = 2.$$

So

$$[2\text{pts}] \quad \mathbb{E}(Y_1Y_2) = \int_0^1 \int_0^{y_2} y_1y_2 g_{12}(y_1, y_2)dy_1dy_2 = 2\int_0^1 \int_0^{y_2} y_1y_2dy_1dy_2$$
$$[1\text{pt}] \quad = 2\int_0^1 y_2 \int_0^{y_2} y_1dy_1dy_2 = 2\int_0^1 y_2\frac{y_2^2}{2}2dy_2 = \frac{1}{4}.$$

Similar to part (a), using (4.4.2) we have the p.d.f. of Y_2 as $g(y_2) = 2y_2$ [2pts], then

$$[1\text{pt}] \quad \mathbb{E}(Y_2) = \int_0^1 2y_2^2 dy_2 = \frac{2}{3}.$$

So

$$[2\text{pts}] \quad Cov(Y_1, Y_2) = \mathbb{E}(Y_1Y_2) - \mathbb{E}(Y_1)\mathbb{E}(Y_2) = \frac{1}{4} - \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) = \frac{1}{36}.$$

4. (20 points) Let $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ denote the sample variance of a random sample $X_1, \ldots, X_n$ from $N(\mu, \sigma^2)$.

   (a) (10 points) Find a $(1-\alpha)$ confidence interval of the format $(0, (n-1)S^2/c)$ for σ^2 using the sampling distribution of S^2 in Student's theorem. Give the constant c as the quantile of a chi-square distribution.

   (b) (10 points) Consider testing $H_0 : \sigma^2 = \sigma_0^2$ v.s. $H_1 : \sigma^2 > \sigma_0^2$. What is the size of the test to reject H_0 if $(n-1)S^2/\sigma_0^2 > c$, where c is the constant in (a)? Explain.

Solution:

(a) [4pts] $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.
[4pts] Let U be a random variable that has a χ^2-distribution with $n-1$ degrees of freedom, and ξ_α is the real number for which $P(U > \xi_\alpha) = 1 - \alpha$. Then

$$P\left[\frac{(n-1)S^2}{\sigma^2} > \xi_\alpha\right] = 1 - \alpha.$$

This is

$$P\left[\sigma^2 < \frac{(n-1)S^2}{\xi_\alpha}\right] = 1 - \alpha.$$

[2pts] Hence, $c = \xi_\alpha$.

(b) [4pts] By definition, size=P(reject H_0 $|H_0$).
[4pts] Under H_0, $\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$, so

[2pts] $$\text{size} = P\left[\frac{(n-1)S^2}{\sigma_0^2} > \xi_\alpha\right] = 1 - \alpha.$$

5. (15 points) The chi-square test statistic for testing the independence between the row and column variables in a 2×2 contingency table is the form

$$Q = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(X_{ij} - O_{ij})^2}{O_{ij}},$$

where X_{ij}'s are the observed cell counts in the table. The following table summarizes results from a study on whether attending class influences how students perform on an exam. For example, 25 students who attended classes passed the exam.

	Pass	Fail
Attended	25	6
Skipped	8	15

(a) (5 points) Suppose that we apply the chi-square test to the above table. Find the values of O_{ij}'s.

(b) (5 points) Under the null hypothesis that the row and column variables are independent, what distribution does the test statistic Q have?

(c) (5 points) The chi-square test is an asymptotic test. Do you see any pitfall in this application of the test? Explain.

Solution:

(a) [1pt]

$$\begin{aligned} N &= 25 + 6 + 8 + 15 = 54 \\ n_{1.} &= 25 + 6 = 31, \;\; n_{.1} = 25 + 8 = 33, \\ n_{2.} &= 8 + 15 = 23, \;\; n_{.2} = 6 + 15 = 21. \end{aligned}$$

[4pts]

$$\begin{aligned} O_{11} &= \frac{n_{1.}n_{.1}}{N} = \frac{31 * 33}{54} = 18.94, \\ O_{21} &= \frac{n_{2.}n_{.1}}{N} = \frac{23 * 33}{54} = 14.06, \\ O_{12} &= \frac{n_{1.}n_{.2}}{N} = \frac{31 * 21}{54} = 12.06, \\ O_{22} &= \frac{n_{2.}n_{.2}}{N} = \frac{23 * 21}{54} = 8.94. \end{aligned}$$

(b) [3pts] The test statistic has a Chi-square distribution.
[2pts] The degree of freedom is

$$df = (2 - 1) * (2 - 1) = 1.$$

(c) [2pts] The chi-square test is sensitive to small expected counts in one or more of the cells in the table.
[3pts] Since expected counts are greater than 5 in this example, there is no problem of applying the chi-square test.

6. (22 points) Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ and a finite variance σ^2.

1. (5 points) Show that $\hat{\theta} = \frac{1}{\bar{X}_n}$ is a consistent estimator of $\theta = \frac{1}{\mu}$.
2. (10 points) Show that $\sqrt{n}(\hat{\theta} - \theta) \to N(0, \frac{\sigma^2}{\mu^4})$ in distribution.
3. (7 points) Explain how to construct an asymptotic $(1 - \alpha)$ confidence interval of θ using the result in (b).

Solution:

(a) [3pts] From the Weak Law of Large Numbers, we know

$$\overline{X}_n \xrightarrow{P} \mu.$$

[2pts] By Theorem 5.1.4, the function $g(x) = \frac{1}{x}$ is continuous at $x = \mu$, we get

$$g(\overline{X}_n) = \frac{1}{\overline{X}_n} \xrightarrow{P} g(\mu) = \frac{1}{\mu}.$$

(b) [4pts] From the Central Limit Theorem, we have

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

[3pts] Using the $\Delta-$method, and $g'(\mu) = -\frac{1}{\mu^2}$,
[3pts]

$$\sqrt{n}(g(\overline{X}_n) - g(\mu)) \xrightarrow{d} N(0, [g'(\mu)]^2\sigma^2) \equiv N(0, \frac{\sigma^2}{\mu^4})$$

(c) [3pts] Based on the distribution in (b), we have

$$P(|\frac{\sqrt{n}(\hat{\theta} - \theta)}{\sigma/\mu^2}| < z_{\alpha/2}) = 1 - \alpha$$

As σ and μ are both unknown, this can not be used directly for constructing a confidence interval. But they can be consistently estimated by $\hat{\sigma} = s$ and $\hat{\mu} = \overline{X}$.
[4pts] By Slutsky's theorem,

$$\frac{\sqrt{n}(\hat{\theta} - \theta)}{s/\overline{X}^2} \xrightarrow{d} N(0, 1)$$

So, $\hat{\theta} \pm z_{\alpha/2}\frac{s}{\overline{X}^2}\frac{1}{\sqrt{n}}$ gives an asymptotic $(1 - \alpha)$ confidence interval of θ.

Practice problems for midterm 1

1. Suppose $X_1, \ldots, X_n$ are a random sample of size n from $Unif(0, \theta)$.

   (a) Show that $2\bar{X}_n$ is an unbiased estimator of θ.

   (b) Let Y_n be the sample maximum, i.e. $\max\{X_1, \ldots, X_n\}$. Find its pdf.

   (c) Find the bias of Y_n as an estimator of θ. And construct an unbiased estimator based on Y_n by correcting its bias.

2. Suppose $X_1, \ldots, X_n$ are a random sample of size n from $Bernoulli(\theta)$, for which the pdf is

   $$f(x; \theta) = \theta^x (1-\theta)^{1-x}, x = 0, 1, 0 < \theta < 1.$$

   Consider testing $H_0 : \theta = 0.5$ versus $H_1 : \theta > 0.5$.

   (a) For a test with critical region $\{(x_1, \ldots, x_n) : \bar{x}_n > c\}$, find the value of c for the test to have an asymptotic significance level $\alpha = 0.05$.

   (b) Suppose $\theta = 0.7$. Derive the power of test in (a).

   (c) If it is observed that $\bar{x}_{25} = 0.7$, should H_0 be rejected at the significance level $\alpha = 0.05$? Make your conclusion based on the p-value from the test in (a).

3. Let $x_1, \ldots, x_n$ be a random sample from $N(\mu, \sigma^2)$, where both parameters μ and σ^2 are unknown. Outline how you construct a 95% confidence interval for σ^2. And what is the expected length of your confidence interval?

4. Suppose $X_1, \ldots, X_n$ are a random sample of size n from $N(0, \theta)$. Then the MLE of θ is $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n X_i^2$.

   (a) Show that $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, 2\theta^2)$ using the Central Limit Theorem.

   (b) Let $g(\theta) = \log\theta$. And show that $\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N(0, c)$, where c is a constant. And give the constant c.

   (c) Find an asymptotic 95% confidence interval for $g(\theta)$.

   (d) Transform the above confidence interval to an asymptotic 95% confidence interval for θ.