Solutions Project 6
14 points

1. (2 pts) The 7 beautiful properties of MLEs are:

1. **1.** An MLE is always a function of a sufficient statistic (see p. 480 of your text for a proof). Many of you in your HW pointed out that the MLE is a sufficient statistic if the likelihood has a unique maximum. This follows because any 1-1 function of a sufficient statistic is a sufficient statistic.
2. **2.** By the Rao-Blackwell Theorem, typically an unbiased function of a sufficient statistic is the MVUE. Hence, Rao-Blackwell suggests that transforming the MLE to an unbiased estimator yields the MVUE.
3. **3.** An MLE is consistent if the likelihood satisfies some regularity conditions.
4. Invariance
5. MLEs are asymptotically efficient if the likelihood satisfies some regularity conditions. That is, Var(MLE) eventually attains the minimum variance possible among all estimators with the same expected value.
6. Var(MLE) attains the Cramer-Rao lower bound (CRLB) among all estimators with the same expected value if the likelihood satisfies some regularity conditions.
7. There is a CLT for MLEs if the likelihood satisfies some regularity conditions. That is, (MLE – parameter)/sqrt(Var(MLE)) converges in distribution to N(0,1).

2. (Exercise 9.98, 2 pts)

To get the numerator of the CRLB estimate of the variance of the MLE, $[d/dp\ t(p)]^2 = [d/dp\ p]^2 = 1^2 = 1$. To get the denominator, consider the natural log of the pmf, $\ln p(y \mid p) = \ln p + (y-1)\ln(1-p)$, so

$$\tfrac{d}{dp}\ln p(y \mid p) = 1/p - (y-1)/(1-p)$$

$$\tfrac{d^2}{dp^2}\ln p(y \mid p) = -1/p^2 - (y-1)/(1-p)^2.$$

Then,

$$-nE\left[\tfrac{d^2}{dp^2}\ln p(Y \mid p)\right] = -E\left[-1/p^2 - (Y-1)/(1-p)^2\right] = \frac{n}{p^2(1-p)}.$$

Therefore, the approximate (limiting) variance of the MLE is given by

$$V(\hat{p}) \approx \frac{p^2(1-p)}{n}.$$

3. (1 pt) By the CLT for MLEs and #2, the MLE $\hat{p}$ is approximately distributed as $N(p, \frac{p^2(1-p)}{n})$ for large n. Hence, an approximate 95% CI for p for large n is $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}^2(1-\hat{p})}{n}}$.

4. (Exercise 9.99, 3pts)

From Example 9.18, the MLE for $t(p) = p$ is $\hat{p} = Y/n$ and the CRLB estimate of Var($\hat{p}$) is equal to the Var($\hat{p}$)=$p(1-p)/n$ that we derived earlier in Chapter 8. Similarly, a 100(1 – α)% CI for p is $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ is the same CI for p derived in Section 8.6.

5. (Exercise 9.100, 3pts)

(a) In Exercise 9.81, it was shown that $\overline{Y}^2$ is the MLE of $t(\theta) = \theta^2$.
(b) The numerator of the asymptotic variance is $[d/d\theta\, t(\theta)]^2 = [d/d\theta\, \theta^2]^2 = [2\theta]^2$. The denominator is

$$-nE\left[\frac{d^2}{dp^2}\ln f(Y\mid\theta)\right] = -n\mathrm{E}\left[\frac{d^2}{d\theta^2}\ln\left(\frac{1}{\theta}e^{-\frac{y}{\theta}}\right)\right] = -n\mathrm{E}\left[\frac{d^2}{d\theta^2}\left(-\ln\theta - \frac{y}{\theta}\right)\right]$$
$$= -n\mathrm{E}\left[\frac{d}{d\theta}\left(-\frac{1}{\theta}+\frac{y}{\theta^2}\right)\right] = -n\mathrm{E}\left[\frac{1}{\theta^2}-\frac{2y}{\theta^3}\right] = -n\left[\frac{1}{\theta^2}-\frac{2\theta}{\theta^3}\right]$$
$$= \frac{n}{\theta^2}.$$

Thus, the asymptotic variance is $4\theta^2/n$.
(c) An approximate large sample 100(1 – α)% CI for θ is

$$\overline{Y}^2 \pm z_{\alpha/2}\sqrt{\left(\frac{(2\theta)^2}{n\frac{1}{\theta^2}}\right)\Bigg|_{\theta=\hat{\theta}}} = \overline{Y}^2 \pm z_{\alpha/2}\left(\frac{2\overline{Y}^2}{\sqrt{n}}\right).$$

6. (Exercise 9.101)

(a) From Exercise 9.80, the MLE for $t(\lambda) = \exp(-\lambda)$ is $t(\hat{\lambda}) = \exp(-\hat{\lambda}) = \exp(-\overline{Y})$.
(b) The numerator of the large sample variance is $(\frac{d}{d\lambda}t(\lambda))^2 = (-\exp(-\lambda))^2 = \exp(-2\lambda)$.
The denominator is

$$-nE\left[\frac{d^2}{dp^2}\ln p(Y\mid\lambda)\right] = -n\mathrm{E}\left[\frac{d^2}{d\theta^2}\ln\left(\frac{\lambda^y e^{-\lambda}}{y!}\right)\right]$$
$$= -n\mathrm{E}\left[\frac{d^2}{d\theta^2}(y\ln\lambda - \lambda - \ln y!)\right] = -n\mathrm{E}\left[\frac{d}{d\theta}(\frac{y}{}/\lambda - 1)\right]$$
$$= -n\mathrm{E}\left[-\frac{y}{\lambda^2}\right] = n\frac{\lambda}{\lambda^2} = n/\lambda.$$

So the large sample variance $\lambda\exp(-2\lambda)/n$.
(c) From part b, a large sample 100(1 – α)% CI for $t(\lambda)=\exp(\lambda)$ is

$$\exp(-\overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\exp(-2\lambda)}{n\frac{1}{\lambda}}}\Bigg|_{\lambda=\overline{Y}} = \exp(-\overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\overline{Y}\exp(-2\overline{Y})}{n}}.$$