SOLUTIONS TO PRACTICE QUESTIONS

17. No, confidence interval means theres a 90% chance the mean is in the confidence interval. It does not represent the probability that an element of the distribution is in the interval.

21. 4 since $var(X/n) = var(X)/n$. Since our CI for $\mathbb{E}[X]$ is of the form $(\mu - sd(X)/n, \mu + sd(X)/n)$ we get that if we quadruple the sample size, the CI is $(\mu - sd(X/n)/2, \mu + sd(X/n)/2)$ which halves the confidence interval.

7. a) This is a negative binomial distribution, so $\mathbb{E}[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \dfrac{1}{p}$ (This is done by breaking the sum up into elements of the form $\sum_{k=n}^{\infty} p(1-p)^{k-1} = (1-p)^{n-1}$

b) $L(p) = \prod_{i=1}^{n} p(1-p)^{k_i - 1}$

So $l(p) = \sum_{i=1}^{n} log(p) + (k_i - 1) log(1-p)$

Taking derivatives gives us $l'(p) = \sum_{i=1}^{n} \dfrac{1}{p} - \dfrac{k_i - 1}{1-p}$

And setting $l'(p) = 0 \Rightarrow \frac{n}{p} = \sum_{i=1}^{n} \dfrac{k_i - 1}{1-p} \Rightarrow p = \dfrac{n}{\sum_{i=1}^{n} k_i}$

c) We note that the second derivative of $f(x)$ is $g(k|p) = -\frac{1}{p^2} - \frac{k-1}{(1-p)^2}$

So $I = -\mathbb{E}[g(k|p)] = \frac{1}{p^2} + \frac{\frac{1}{p}-1}{(1-p)^2} = \frac{1}{p^2(1-p)}$

Therefore, the asymptotic variance of the MLE is $\frac{1}{nI} = \frac{p^2(1-p)}{n}$

17. a) The parameter α determines how concentrated the distribution is at $\frac{1}{2}$. This can be seen through the symmetry of the distribution as well as the fact that $x(1-x)$ is maximized at $\frac{1}{2}$.

b) We set $var(X) = \sum_{i=1}^{n} \dfrac{X_i^2}{n} - \dfrac{1}{4}$, the sample variance

Then we get that $\frac{1}{4(2\alpha+1)} = \sum_{i=1}^{n} \dfrac{X_i^2}{n} - \dfrac{1}{4}$

By solving for α, we get that $\alpha = \frac{1}{2}\left(\frac{1}{4(\sum_{i=1}^{n} \frac{X_i^2}{n} - \frac{1}{4})} - 1\right) = \frac{n}{4\sum_{i=1}^{n} X_i^2 - 2n} - \frac{1}{2}$

c) $\frac{d}{d\alpha} log(\prod_{i=1}^{n} f(x_i|\alpha)) = \dfrac{d}{d\alpha} \sum_{i=1}^{n} (log(\Gamma(2\alpha)) - 2log(\Gamma(\alpha)) + (\alpha - 1) log(x_i(1-x_i)))$

$= \dfrac{2n\Gamma'(2\alpha)}{\Gamma(2\alpha)} - \dfrac{2n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n} log(x_i(1-x_i))$

Setting equal to 0 and dividing by 2n gives solution in book

d) $log(f(x|\alpha)) = log(\Gamma(2\alpha)) - 2log(\Gamma(\alpha)) + (\alpha - 1) log(x_(1 - x_))$

Taking 2 derivatives with respect to α yields the solution in the manual (using quotient rule for second derivative)

e) Taking $T(x_1,...,x_n,\theta) = \prod_{i=1}^{n} x_i(1-x_i)$ we see that the joint density of $x_1,..,x_n$ is

$\prod_{i=1}^{n} f(x_i|\alpha) = \prod_{i=1}^{n} \dfrac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} x_i(1-x_i) = (\dfrac{\Gamma(2\alpha)}{\Gamma(\alpha)^2})^n T(x_1,...,x_n,\theta)$

By the factorization theorem, T is a sufficient statistic. (taking h=1, $g(T,\alpha) = (\frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2})^n T(x_1,...,x_n,\theta)$

19. c) no, MLE are minimum variance unbiased estimators.

47. a) $\mathbb{E}[X] = \frac{\theta}{\theta-1} x_0$

Therefore, the method of moments estimate is $\theta = \frac{\mathbb{E}[X]}{\mathbb{E}[X]-x_0}$ where we replace $\mathbb{E}[X]$ with the sample mean $\overline{X}$

b) $\prod_{i=1}^{n} f(x_i) = \theta^n x_0^{\theta n} (\prod_{i=1}^{n} x_i)^{-\theta-1}$

Taking derivatives with respect to theta and setting equal to 0 yields

$0 = \frac{\theta}{n} + nlog(x_0) - \sum_{i=1}^{n} log(x_i)$

So $\theta = \dfrac{n}{\sum_{i=1}^{n} log(x_i) - nlog(x_0)}$

c) taking 2 derivatives of $log(f(x|\theta))$ we get $-\frac{1}{\theta^2}$

Therefore, $I = \frac{1}{\theta^2}$ so that the asymptotic variance is $\frac{\theta^2}{n}$

d) The joint density is $\prod_{i=1}^{n} \theta x_0^{\theta} x_i^{-\theta-1} = (\theta x_0^{\theta})^n (\prod_{i=1}^{n} x_i)^{-\theta-1}$

So letting $T = \prod_{i=1}^{n} x_i$, this is a function of T and θ

Therefore, by the factorization theorem, taking $g(T,\theta) = (\theta x_0^{\theta})^n (\prod_{i=1}^{n} x_i)^{-\theta-1}, h = 1$ we get that $\prod_{i=1}^{n} x_i$ is a sufficient statistic.

59. a) We get that $P(MF) = (1-\alpha)\frac{1}{2}$ since there is a $1-\alpha$ chance of not being identical, and a $\frac{1}{2}$ of being MF if they are independent

Therefore, $P(MM) + P(FF) = 1 - P(MF) = \frac{1+\alpha}{2}$. Since $P(MM) = P(FF)$ we get $P(MM) = P(FF) = \frac{1+\alpha}{4}$

b) $L(\alpha) = (\frac{1+\alpha}{4})^{n_1+n_2}((1-\alpha)\frac{1}{2})^{n_3}$

so $l(\alpha) = (n_1+n_2)log(\frac{1+\alpha}{4}) + n_3 log((1-\alpha)\frac{1}{2})$

Taking derivatives and setting equal to 0 yields $(n_1+n_2)/(1+\alpha) - n_3/(1-\alpha) = 0$

This means that $(n_1+n_2)(1-\alpha) - n_3(1+\alpha) = 0$ so that $\alpha = \frac{n_1+n_2-n_3}{n_1+n_2+n_3}$

Although technically this is always true, in order for it to make sense in the context of the problem, α must be positive.

Taking 2 derivatives of $log(f(x|\alpha))$ we get $-\frac{1}{(1-\alpha)^2}$ for MF, and $-\frac{1}{(1+\alpha)^2}$ for MM, FF

So $I = \frac{1}{(1-\alpha)^2} P(MF) + 2\frac{1}{(1+\alpha)^2} P(MM) = \frac{1}{2(1-\alpha)} + \frac{1}{2(1+\alpha)} = \frac{\alpha}{(1-\alpha)(1+\alpha)}$
Therefore, the asymptotic variance is $\frac{(1-\alpha)(1+\alpha)}{n\alpha}$

69. The joint density of $(x_1, .., x_n)$ for a geometric distribution is

$$\prod_{i=1}^{n} p(1-p)^{x_i - 1} = p^n (1-p)^{\sum_{i=1}^{n} x_i - n} = p^n (1-p)^{T-n}$$

Therefore, since the joint distribution can be factored into

$g(T, p) = p^n (1-p)^{\sum_{i=1}^{n} x_i - n} = p^n (1-p)^{T-n}, h = 1$ we have that T is a sufficient statistic