Stat 135, Fall 2006 **Prof. Adhikari**

Midterm Answers

1.

Estimate	**Unbiased?**	**MOM?**	**MLE?**
$\bar{X}$ as an estimate of μ	yes	yes	yes
$\hat{\sigma}^2$ as an estimate of σ^2	no	yes	yes
S^2 as an estimate of σ^2	yes	no	no

2.

the object	**real/r.v.**
the 95th percentile of the t distribution with 4 degrees of freedom	real
a generalized likelihood ratio	r.v.
the variance of the MLE of a parameter	real
the p-value of a test	r.v.

3a. The likelihood is

$$\frac{1}{\theta} \cdot \frac{1}{\theta} \cdot \frac{\theta - 6}{\theta}, \qquad \theta \ > \ 6$$

b. Simplify the likelihood function: it is

$$\frac{1}{\theta^2} - \frac{6}{\theta^3}$$

You can take logs if you like, but the function is easy enough to differentiate directly. Its derivative with respect to θ is

$$\frac{-2}{\theta^3} + \frac{18}{\theta^4}$$

Set this equal to 0 and solve:

$$-2 \ + \ \frac{18}{\hat{\theta}} \ = \ 0 \quad \text{so} \quad \hat{\theta} \ = \ 9.$$

4a. Small sample, normal population with unknown SD. The interval is $2.77 \pm t0.61/\sqrt{15}$ where t is the 97.5th percentile of the t distribution with 14 degrees of freedom.

b. Treat each mouse as a coin toss, and let p be the probability that the mouse runs faster after treatment. Then test $H_0 : p = 0.5$ versus $H_0 : p > 0.5$, using the binomial distribution with $n = 15$ and $p = 0.5$ to find the p-value. In the sample, 11 out of the 15 mice ran faster the second time. So p is the probability of getting 11 or more heads in 15 tosses of a fair coin.

You can also use the differences in times and test whether the underlying mean of the differences (pre minus post) is equal to 0 or greater than 0. You would use the t curve with 14 d.f.

There's nothing to pool here. You've got one sample of pairs, not two independent samples.

5a. The Neyman Pearson lemma says that the likelihood ratio test is the most powerful. The ratio is

$$\frac{10e^{-10X}}{20e^{-20X}} \;=\; 0.5e^{10X}$$

which is small when X is small. (You can guess that anyway. Under H_0 the expectation of X is $1/10$ and under H_A it's $1/20$, so small values of X favor the alternative.)

If the critical value is called c, then

$$0.1 \;=\; P_{H_0}(X < c) \;=\; 1 - e^{-10c}; \text{ so } e^{-10c} \;=\; 0.9; \text{ so } -10c \;=\; \log 0.9; \text{ so } c \;=\; \frac{\log 0.9}{-10}$$

b. The power is

$$P_{H_A}(X < c) \;=\; 1 - e^{-20c} \;=\; 1 - e^{2\log 0.9} \;=\; 1 - 0.9^2 \;=\; 0.19$$

6a. T^2 is unbiased because

$$E(T^2) \;=\; E\Big[\frac{1}{n}\sum_{i=1}^{n}(X_i - 10)^2\Big] \;=\; \frac{1}{n}\sum_{i=1}^{n}E(X_i - 10)^2 \;=\; E(X_1 - 10)^2 \;=\; \sigma^2$$

by the definition of the variance of X_1.

b. (For some reason this one is labeled **c** on the test.) A chi-squared variable with n degrees of freedom is the of sum squares of n i.i.d. standard normals. So

$$Y = \sum_{i=1}^{n}\Big(\frac{X_i - 10}{\sigma}\Big)^2$$

has the chi-squared distribution with n degrees of freedom. Clearly $Y = (n\sigma^2)T$, so the constants are $c = n/\sigma^2$ and $d = n$.

c. (Called **d** in the test.) By the previous part, $T = (1/c)Y$ so $Var(T) = Var(Y)/c^2$. The variance of a chi-squared random variable with n degrees of freedom is $2n$, so

$$Var(T) \;=\; 2n \cdot \frac{\sigma^4}{n^2} \;=\; \frac{2\sigma^4}{n}$$

If you didn't remember the variance of the chi-squared, you could have used the fact that the chi-squared is a particular gamma, and then plug into the formula for the gamma variance which was given in the problem.

d. (a.k.a. **e**). You have to check that $Var(T^2)$ attains the Cramer-Rao bound. Rename σ^2 to be θ, as suggested in the problem. The normal density at X is

$$f(X|\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{\frac{(X-10)^2}{2\theta}}$$

and so

$$\log f(X|\theta) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\theta - \frac{(X-10)^2}{2\theta}$$

Differentiate once with respect to θ:

$$-\frac{1}{2\theta} + \frac{(X-10)^2}{2\theta^2}$$

Differentiate again:

$$\frac{1}{2\theta^2} - \frac{(X-10)^2}{\theta^3}$$

Remember that $E[(X-10)^2] = \sigma^2 = \theta$ to see that the Fisher information is

$$I(\theta) = -E\Big[\frac{1}{2\theta^2} - \frac{(X-2)^2}{\theta^3}\Big] = -\frac{1}{2\theta^2} + \frac{\theta}{\theta^3} = \frac{1}{2\theta^2} = \frac{1}{2\sigma^4}$$

Therefore according to the Cramer-Rao bound, the variance of any unbiased estimate of σ^2 is at most $2\sigma^4/n$, which is exactly the variance of T^2.

So T^2 is indeed the most efficient among all unbiased estimates of σ^2.