Name: ____________________________

Directions: **Work only on this sheet** (on both sides, if needed); do not turn in any supplementary sheets of paper. There is actually plenty of room for your answers, as long as you organize yourself BEFORE starting writing. In order to get full credit, **SHOW YOUR WORK**.

NOTE: ALL PROBLEMS HERE REQUIRE CLOSED-FORM SOLUTIONS.

1. Suppose $f_W(t) = ct^{c-1}$ for t in (0,1), with the density being 0 elsewhere, for some unknown $c > 0$. We have a random sample $W_1, ..., W_n$ from this density.

(a) (10) Find F_W.

(b) (15) Find the Method of Moments Estimator of c. It should work out to be a function of $\bar{W}$.

(c) (15) Suppose c = 2 and n = 50. Find the value of d for which $F_{\bar{W}}(d) \approx 0.975$.

2. Suppose $f_W(t) = \frac{1}{c}$ on (0,c), 0 elsewhere, for some $c > 0$. We have a random sample $W_1, ..., W_n$ from this density.

(a) (15) Find the Maximum Likelihood Estimator (MLE).

(b) (15) Find the bias of the MLE.

3. (15) Suppose the Y = 1 or 0, with probability 0.6 and 0.4. If Y = 1, X has the density $3t^2$ on (0,1); if Y = 0, then X's density is 2t. Find $m_{Y;X}(t)$.

4. (15) Suppose we assume that $Var(Y|X = t) = \sigma^2$, independent of t, but do not assume linearity. Which of the statements below is true? **BE SURE TO JUSTIFY YOUR ANSWER, USING MATH.**

(i) $Var(Y) \leq \sigma^2$

(ii) $Var(Y) \geq \sigma^2$

(iii) Either (i) or (ii) could be true, depending on the specific distributions.

Solutions:

1.a For t in (0,1),

$$F_W(t) = P(W \leq t) = \int_0^t cs^{c-1}\,ds = t^c \tag{1}$$

$F_W(t)$ is 0 for t less than 0, and is 1 for t bigger than 1.

1.b

$$EW = \int_0^1 tct^{c-1}\,dt = \frac{c}{c+1} \tag{2}$$

So, set

$$\bar{W} = \frac{\hat{c}}{\hat{c}+1} \tag{3}$$

yielding

$$\hat{c} = \frac{\bar{W}}{1-\bar{W}} \tag{4}$$

1.c By the Central Limit Theorem $\bar{W}$ is approximately normally distributed with mean EW and variance EW/n^2.

We already found that EW = c/(c+1) = 2/3 in this case. Also $E(W^2)$ turns out to be 1/2, so Var(W) = 1/18. Thus

$$Var(\hat{W}) = \frac{1/18}{50} \tag{5}$$

By the CLT, the 97.5th percentile is about 1.96 standard deviations above the mean, i.e.

$$d \approx \frac{2}{3} + 1.96\sqrt{\frac{1}{900}} \approx \frac{2}{3} + \frac{1}{15} = 11/15 \tag{6}$$

NOTE THAT YOU DID NOT NEED A CALCULATOR!

2.a The likelihood is

$$\left(\frac{1}{c}\right)^n \tag{7}$$

as long as

$$c \geq max_i W_i \tag{8}$$

So,

$$\hat{c} = max_i W_i \tag{9}$$

just as in our discrete example in the PLN (conference paper numbers).

2.b The bias is $E\hat{C} - c$. To get $E\hat{c}$ we need the density of that estimator, which we get as follows:

$$P(\hat{c} \leq t) = P(\text{all } W_i \leq t) = \left(\frac{t}{c}\right)^n \tag{10}$$

So,

$$f_{\hat{c}}(t) = \frac{n}{c^n} t^{n-1} \tag{11}$$

Integrating against t, we find that

$$E\hat{C} = \frac{n}{n+1} c \tag{12}$$

So the bias is c/(n+1), not bad at all.

3 If Y takes on only the values 0 and 1, then $m_{Y;X}(t) = P(Y=1|X=t)$. From our PLN, the latter quantity is

$$\frac{1}{1+\frac{(1-q)f_{X|Y=0}(t)}{qf_{X|Y=1}(t)}} \tag{13}$$

The information given in the problem implies that this last expression is

$$\frac{1}{1+\frac{0.4}{0.6}\frac{2t}{3t^2}} = \frac{1}{1+\frac{4}{9t}} \tag{14}$$

4. From the "law of total variance" in our PLN,

$$Var(Y) = E[Var(Y|X)] + Var[E(Y|X)] \tag{15}$$

That first term on the right is

$$E[\sigma^2] = \sigma^2 \tag{16}$$

The second term is a variance, and is thus nonnegative. Therefor

$$Var(Y) \geq \sigma^2 \tag{17}$$