Math 404 Exam 2 Spring 2016 Name____________________

YOU ARE BEING GRADED FOR WORK

1) Suppose a Poisson (θ) distribution is the fitted distribution where $\overline{X} = 2$. Fill in the following table that would be used for a χ^2 test.

interval	$O_i = n_i$	p_i	$E_i = np_i$	$C_i = \frac{(O_i - E_i)^2}{E_i}$
0	9	0.1353	13.53	1.5167
1	34	0.2707	27.07	1.7741
2	22	.2707	27.07	0.9496
3	23	.1804	18.04	1.3637
4+	$\frac{12}{100}$	.1429	14.29	0.3670

$\sum E_i = n$

$\sum O_i = n = 100$

$$P_i = \frac{e^{-\theta}\theta^i}{i!}$$

$$\hat{\theta} = 2$$

$$\hat{p}_0 = e^{-2}, \quad \hat{p}_1 = 2e^{-2}, \quad \hat{p}_2 = \frac{2^2 e^{-2}}{2} = 2e^{-2}$$

$$\hat{p}_3 = \frac{2^3 e^{-2}}{6} = \frac{4}{3}e^{-2} = .1804$$

10

$$\hat{p}_{4+} = P(X \geq 4) = 1 - \hat{p}_0 - \hat{p}_1 - \hat{p}_2 - \hat{p}_3$$

2) You observe the following 4 ground up claims ← so x_i from a data set that is truncated below at 1000 (no losses below 1000 will be submitted to the insurance company).
1028.5334, 1445.7569, 1952.3840, 3529.0490

Find the MLE of θ if an EXP(θ) distribution is fit to the data.

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - d) = \frac{28.5334 + 445.7569 + 952.384 + 2529.049}{4}$$

$$= \frac{3955.7233}{4} = \boxed{988.9308}$$

10

(y^P are payments x are policyholder losses)

Y^P are payments
X are policyholder losses

3) An insurance policy has a deductible of 20 and a policy limit (maximum payment per loss) of 80. The three insurer loss payments under this policy have been 10, 30, 50, and 80. The insurer is only aware of those losses on which a payment is made. Assume that the ground–up losses were generated by an exponential distribution with a mean of θ.

a) You are given Y^p. Convert the Y^P values to X values that have been truncated and possibly censored using $X = Y^P + d$.

Y^P 10 30 50 80 add 20

$\boxed{X \quad 30 \quad 50 \quad 70 \quad 100}$ ← censored

X is truncated at $d = 20$ censored at $u = 100$

b) Find $L(\theta)$.

$$\frac{f(30)\, f(50)\, f(70)\, [1-F(100)]}{[1-F(20)]^4}$$ ← all 4 are truncated

$$= \frac{\frac{1}{\theta} e^{-\frac{30}{\theta}} \frac{1}{\theta} e^{-\frac{50}{\theta}} \frac{1}{\theta} e^{-\frac{70}{\theta}} e^{-\frac{100}{\theta}}}{\left[e^{-\frac{20}{\theta}}\right]^4}$$

$$= \frac{1}{\theta^3} e^{-(30+50+70+100-80)} = \boxed{\frac{1}{\theta^3} e^{-170/\theta}}$$

c) Find the MLE of θ.

$$\ln L(\theta) = -3\ln\theta - \frac{170}{\theta}$$

$$\frac{d \ln L(\theta)}{d\theta} = -\frac{3}{\theta} + \frac{170}{\theta^2} \overset{\text{set}}{=} 0 \quad \text{or } 170 = 3\hat{\theta}$$

$$\hat{\theta} = \frac{170}{3} = \boxed{56.6667}$$

$$\text{or } \hat{\theta} = \frac{\sum (x_i - d)}{m} = \frac{\sum Y_i^P}{m} = \frac{10+30+50+80}{3} = \frac{170}{3} = \boxed{56.6667}$$

easy way

uncensored

$$\tau(\theta) = \theta^{\frac{1}{2}}, \quad \tau'(\theta) = \frac{1}{2}\theta^{-\frac{1}{2}} = \frac{1}{2\sqrt{\theta}} \quad (\tau'(\theta))^2 = \frac{1}{4\theta}$$

4) Suppose $X_1, ..., X_n$ are iid with $I_1(\theta) = \dfrac{2}{\theta^2}$ where $\theta > 0$. Find the asymptotic variance of the MLE of $\sqrt{\theta}$.

$$var(\tau(\hat{\theta})) = [\tau'(\theta)]^2 \, var(\hat{\theta}) =$$

$$[\tau'(\theta)]^2 \underbrace{\frac{1}{nI_1(\theta)}} = \frac{1}{4\theta \, n \frac{2}{\theta^2}} = \boxed{\frac{\theta}{8n}}$$

$$\frac{\hat{\theta}}{8n} \text{ ok}$$

$$\left(var(\hat{\theta}) = \frac{1}{nI_1(\theta)}\right)$$

10

5) Let $f(x) = (\theta + 1)x^\theta$ be a one parameter exponential family where $\theta > -1$ and $0 < x < 1$. Find $I_1(\theta)$.

$$\left(f(x) = (\theta+1)\exp[\theta(\ln x)]\right)$$

$$\ln f(x) = \ln(\theta+1) + \theta \ln(x)$$

$$\frac{d}{d\theta} \ln f(x) = \frac{1}{\theta+1} + \ln x$$

$$\frac{d^2}{d\theta^2} \ln f(x) = \frac{-1}{(\theta+1)^2} \quad = \frac{d}{d\theta}(\theta+1)^{-1}$$

$$I_1(\theta) = -E\left(\frac{-1}{(\theta+1)^2}\right) = \boxed{\frac{1}{(\theta+1)^2}}$$

10

e

6) Suppose the fitted distribution gave the following table. Fill in the table and do a 4 step χ^2 test.

type	$O_i = n_i$	p_i	$E_i = np_i$	$C_i = \frac{(O_i - E_i)^2}{E_i}$
A	86	0.5	100	1.96
B	45	0.2	40	0.625
C	69	.3	60	1.35
	200	1	200	3.935

$\sum O_i = \sum E_i = 200$

$\sum p_i = 1$

i) H_0 fitted dist is good H_1 / not H_0

ii) $Q = 3.935$

iii) $df = k - r - 1 = 3 - 0 - 1 = 2$ | $.95$: 5.991

Use $\alpha = .05$ if not given α

$Q < 5.991$ fail to reject H_0

iv) the fitted dist is good (or not enough evidence to say fitted dist is not good)

10e

7) Decide whether a Poisson, Negative Binomial or Binomial distribution best fits the data given in the table below.

number of accidents	number of policies	$\frac{kn_k}{n_{k-1}}$
0	425	
1	401	$401/425 = .9435$
2	154	$2(154)/401 = .7681$
3	19	$3(19)/154 = .3701$
4	1	($4(1)/19 = .2105$ too small)
total	1000	

$\bar{X} > \hat{\sigma}^2$

$\boxed{\text{binomial}}$

$\boxed{\text{decreasing so binomial}}$

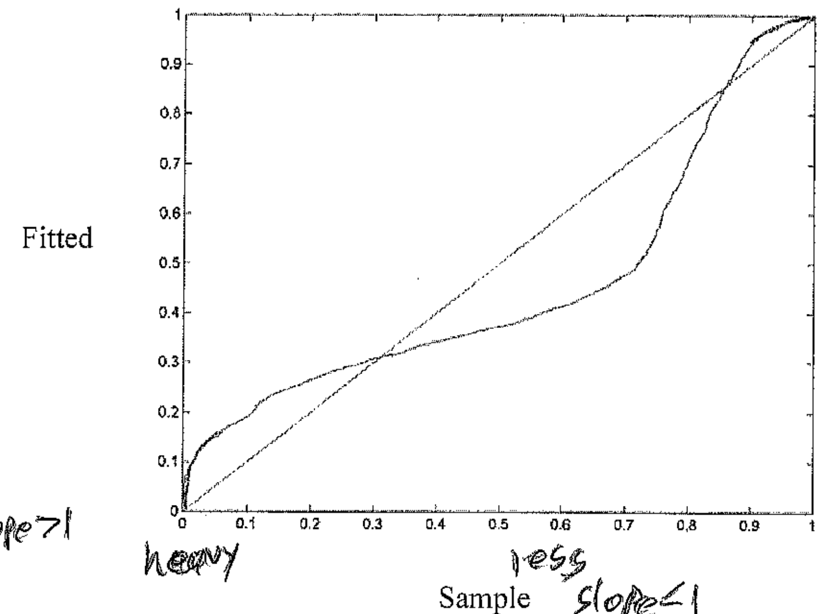
or $\bar{X} = \frac{401 + 154(2) + 19(3) + 1(4)}{1000} = \frac{770}{1000} = 0.77$

$\frac{1}{n}\sum x_i^2 = \frac{401 + 154(2^2) + 19(3^2) + 1(4^2)}{1000} = \frac{1204}{1000} = 1.204$

10

$\hat{\sigma}_E^2 = t - m^2 = 1.204 - (.77)^2 = 0.6111$ $\left(\hat{\sigma}_U^2 = \frac{1000}{999}\hat{\sigma}_E^2 = 0.6118\right)$

59. The graph below shows a *p-p* plot of a fitted distribution compared to a sample.



light slope<1

slope>1 heavy

less slope<1

Which of the following is true?

(A) The tails of the fitted distribution are too thick on the left and on the right, and the fitted distribution has less probability around the median than the sample.

(B) The tails of the fitted distribution are too thick on the left and on the right, and the fitted distribution has more probability around the median than the sample.

(C) The tails of the fitted distribution are too thin on the left and on the right, and the fitted distribution has less probability around the median than the sample.

(D) The tails of the fitted distribution are too thin on the left and on the right, and the fitted distribution has more probability around the median than the sample.

(E) The tail of the fitted distribution is too thick on the left, too thin on the right, and the fitted distribution has less probability around the median than the sample.

10