Final
Stat 490C
December 20, 2008

1. (5 points) A sample from a binomial distribution has a mean of 1.65 and a variance of 0.66.

   Determine $\hat{m}$ and $\hat{q}$ using the Method of Moment Matching.

$$\mu = \hat{m}\hat{q} = 1.65$$
$$Var = \hat{m}\hat{q}(1-\hat{q}) = .66$$
$$\frac{\hat{m}\hat{q}(1-\hat{q})}{\hat{m}\hat{q}} = \frac{.66}{1.65}$$

$$1-\hat{q} = .4$$
$$\hat{q} = .6$$
$$\hat{m} = \frac{1.65}{.6} = 2.75$$

But m must be an integer so

$$\hat{m} = 3$$
$$\hat{q} = \frac{1.65}{3} = .55$$

2. (7 points) You are given the following random sample from an Exponential distribution:

3 6 8 10 20

Using percentile matching at the 20th percentile, estimate θ.

$n = 5 \qquad (n+1)(.2) = (6)(.2) = 1.2$

$j = 1$

$h = .2$

$\pi_{.2} = (3)(1-.2) + (6)(.2) = 2.4 + 1.2 = 3.6$

$F(3.6) = .2 = 1 - e^{-\frac{3.6}{\theta}}$

$e^{-\frac{3.6}{\theta}} = .8$

$-\frac{3.6}{\theta} = \ln(.8)$

$\theta = \frac{-3.6}{\ln(.8)} = 16.133$

3. (7 points) Lee Life Insurance Company is completing a mortality study on a 4 year term insurance policy. The following data is available:

Life	Date of Entry	Date of Exit	Reason for Exit
1	0	0.2	Lapse
2	0	0.3	Lapse
3	0	0.5	Lapse
4	0	0.5	Death
5	0	0.7	Lapse
6	0	1.0	Death
7	0	2.0	Lapse
8	0	2.5	Death
9	0	3.0	Lapse
10	0	3.5	Death
11	0	4.0	Expiry of Policy
12	0	4.0	Expiry of Policy
13	0	4.0	Expiry of Policy
14	0	4.0	Expiry of Policy
15	0	4.0	Expiry of Policy
16	0	4.0	Expiry of Policy
17	0	4.0	Expiry of Policy
18	0	4.0	Expiry of Policy
19	0.5	4.0	Expiry of Policy
20	0.7	1.0	Death
21	1.0	3.0	Death
22	1.0	4.0	Expiry of Policy
23	2.0	2.5	Death
24	2.0	2.5	Lapse
25	3.0	3.5	Death

The Company's chief actuary, Sophie, decides to calculate $S_{25}(2.5)$ using the Kaplan-Meier product-limit estimator where death is the decrement of interest. What did she get for $S_{25}(2.5)$.

j	y_j	s_j	r_j
1	.5	1	16
2	1.0	2	15
3	2.5	2	16

$$S_{25}(2.5) = \frac{15}{16} \cdot \frac{13}{15} \cdot \frac{14}{16} = .709375$$

4. (7 points) Lee Life Insurance Company is completing a mortality study on a 4 year term insurance policy. The following data is available:

Life	Date of Entry	Date of Exit	Reason for Exit
1	0	0.2	Lapse
2	0	0.3	Lapse
3	0	0.5	Lapse
4	0	0.5	Death
5	0	0.7	Lapse
6	0	1.0	Death
7	0	2.0	Lapse
8	0	2.5	Death
9	0	3.0	Lapse
10	0	3.5	Death
11	0	4.0	Expiry of Policy
12	0	4.0	Expiry of Policy
13	0	4.0	Expiry of Policy
14	0	4.0	Expiry of Policy
15	0	4.0	Expiry of Policy
16	0	4.0	Expiry of Policy
17	0	4.0	Expiry of Policy
18	0	4.0	Expiry of Policy
19	0.5	4.0	Expiry of Policy
20	0.7	1.0	Death
21	1.0	3.0	Death
22	1.0	4.0	Expiry of Policy
23	2.0	2.5	Death
24	2.0	2.5	Lapse
25	3.0	3.5	Death

Calculate the Var[$S_{25}(2.5)$] using the Greenwood approximation where death is the decrement of interest.

$$\text{Var}\left[S_{25}(2.5)\right] = \left[S_{25}(2.5)\right]^2 \sum_{i=1}^{3} \frac{s_i}{(r_i)(r_i - s_i)}$$

Using Data From 3

$$(.7109375)^2 \left[\frac{1}{(16)(15)} + \frac{2}{(15)(13)} + \frac{2}{(16)(14)} \right]$$

$$= .011802673$$

5. During a one-year period, the amount paid on accidents incurred by Allen Auto Insurance is:

Amount of Claim	Count
0 – 1,000	20
1,000 – 2,500	25
2,500 – 5,000	30
5000 – 10,000	20
10,000 +	5

H_0: Claim amounts are distributed as Paredo with $\theta = 10000$ and $\alpha = 3$.

H_1: Claim amounts are not distributed as Paredo with $\theta = 10000$ and $\alpha = 3$.

(6 points) Calculate the chi-square statistic.

$$\chi^2 = \sum \frac{(E_j - O_j)^2}{E_j} \qquad E_j = n\left[F(c_j) - F(c_{j-1})\right]$$

$$= 100\left[\left(\frac{10000}{10000 + c_{j-1}}\right)^3 - \left(\frac{10000}{10000 + c_j}\right)^3\right]$$

	O_j	E_j	$\frac{(E_j - O_j)^2}{E_j}$
0 – 1000	20	$100\left[\frac{10000}{10000}\right]^3 - 100\left[\frac{10000}{11000}\right]^3 = 24.8085$	0.9531
1000 – 2500	25	$100\left[\left(\frac{10000}{11000}\right)^3 - \left(\frac{10000}{12500}\right)^3\right] = 23.9315$	0.0477
2500 – 5000	30	$100\left[\left(\frac{10000}{12500}\right)^3 - \left(\frac{10000}{15000}\right)^3\right] = 21.5704$	3.2943
5000 – 10000	20	$100\left[\left(\frac{10000}{15000}\right)^3 - \left(\frac{10000}{20000}\right)^3\right] = 17.1296$	0.4810
10000 +	5	$100\left[\left(\frac{10000}{20000}\right)^3 - 0\right] = 12.5000$	4.5000
	100		9.2761

$= \chi^2$

(2 points) Calculate the critical value at a 5% significance level.

degrees of freedom = 5 − 1

critical value = 9.488

(1 point) State whether you would reject the H_0 at a 5% significance level.

Since $\chi^2 < 9.488$

you should not reject H_0

6. Five iPhone's are observed until they fail with failure measured in months. The iPhones have the following failure dates:

   8 12 15 25 40

   (3 points) If the data is smoothed using a triangular kernel with a bandwidth of 6, calculate variance of the smoothed distribution.

$$Var = Var_e + \frac{b^2}{6}$$

$$E(x) = \frac{100}{5} = 20$$

$$E[x^2] = \frac{64 + 144 + 225 + 625 + 1600}{5} = 531.6$$

$$Var_e = 531.6 - (20)^2 = 131.6$$

$$Var = 131.6 + \frac{6^2}{6} = 137.6$$

   (3 points) If the data is smoothed using a gamma kernel with a bandwidth of 6, calculate the variance of the smoothed distribution.

$$Var = Var_e + \frac{E[x^2]}{\alpha}$$

$$= 131.6 + \frac{531.6}{6} = 220.20$$

7. During a one-year period, Brian Ball Bearing Factory suffered the following workers compensation losses:

300 500 1400 3800

The losses are assumed to be distributed following an Exponential distribution with θ estimated by the maximum likelihood estimator.

Irene wants to test this hypothesis using the Kolmogorov-Smirnov test with a significance level of 1%.

(6 points) Calculate the Kolmogorov-Smirnov test statistic.

$\bar{X} = \hat{\theta} = 1500$ $\quad F^*(x) = 1 - e^{-\frac{x}{1500}}$

	$F_n(x^-)$	$F_n(x)$	$F^*(x)$	D^*
300	0	1/4	0.181269	0.181269
500	1/4	1/2	0.283469	0.216531
1400	1/2	3/4	0.606759	0.143241
3800	3/4	1	0.920606	0.170606

Max = 0.216531 $= D^*$

(1 point) State the critical value for the test.

$$\frac{1.63}{\sqrt{n}} = \frac{1.63}{2} = .815$$

(1 point) State your conclusion.

D < critical value

0.216531 < 0.815 Do Not Reject

(1 point) Would your conclusion be different at the 10% significance level?

critical value $= \frac{1.22}{\sqrt{n}} = \frac{1.22}{2} = .61$

STILL DO NOT REJECT

8. (6 points) You are given the following random sample from a Gamma distribution:

5 10 24

Using the method of moments, estimate α and θ.

$$\bar{X} = \frac{5+10+24}{3} = 13$$

$$Var = \frac{(5-13)^2 + (10-13)^2 + (24-13)^2}{3}$$

$$= 64.6\bar{6}$$

$$\mu = \alpha\theta$$

$$Var(X) = \alpha\theta^2$$

$$\frac{Var}{\mu} = \theta = \frac{64.6\bar{6}}{13} = 4.974358974$$

$$\alpha = \frac{\mu}{\theta} = \frac{13}{4.974358974} = 2.613402$$

9. (10 points) An insurance policy pays claims up to a limit of 2000. A random sample of three payments is obtained as follows: 300, 1000, and 2000.

The claims are assumed to follow a Pareto distribution with $\theta = 2000$.

Calculate the maximum likelihood estimate for α.

$$L(\theta) = f(300)\, f(1000)\, [1 - F(1000)]$$

$$= \frac{\alpha (2000)^{\alpha}}{(300+2000)^{\alpha+1}} \cdot \frac{\alpha (2000)^{\alpha}}{(1000+2000)^{\alpha+1}} \cdot \left(\frac{2000}{2000+2000}\right)^{\alpha}$$

$$= \frac{\alpha}{2300}\left(\frac{2000}{2300}\right)^{\alpha}\left(\frac{\alpha}{3000}\right)\left(\frac{2000}{3000}\right)^{\alpha}\left(\frac{2000}{4000}\right)^{\alpha}$$

$$= \frac{\alpha^2}{(2300)(3000)} \cdot \left(\frac{20 \cdot 2 \cdot 1}{23 \cdot 3 \cdot 2}\right)^{\alpha} = \frac{\alpha^2}{(2300)(3000)} \cdot \left(\frac{40}{138}\right)^{\alpha}$$

$$= \frac{\alpha^2}{(2300)(3000)} \cdot \left(\frac{20}{69}\right)^{\alpha}$$

$$l(\theta) = 2 \ln(\alpha) - \ln[(2300)(3000)] + \alpha \ln 20 - \alpha \ln 69$$

$$\frac{d}{d\theta} \ln(\theta) = \frac{2}{\alpha} - 0 + \ln 20 - \ln 69 = 0$$

$$\alpha = \frac{2}{\ln 69 - \ln 20} = 1.615021$$

10. (4 points) An insurance policy pays claims up to a limit of 2000. A random sample of three payments is obtained as follows: 300, 1000, and 2000.

The claims are assumed to follow an Exponential distribution.

Calculate the maximum likelihood estimate for θ.

$$\hat{\theta} = \frac{\text{TOTAL PAYMENTS}}{\text{\# of Uncensored Payments}}$$

$$= \frac{300 + 1000 + 2000}{2}$$

$$= 1650$$

11. (4 points) An insurance policy pays claims up to a limit of 2000. A random sample of three claim **payments** is obtained as follows: 300, 1000, and 2000.

The claims **amounts** are assumed to follow a Uniform distribution on (0, ω).

Based on the data, we can surmise that the claim **amounts** could be grouped as follows:

Claim Amounts	Count
0 – 2000	2
2000 +	1

Calculate the maximum likelihood estimate for ω.

$$\hat{\omega} = \frac{\#\text{ of claims}}{\text{claims} < \text{largest breakpoint}} \times \text{largest breakpoint}$$

$$= \frac{3}{2} \cdot 2000 = 3000$$

12. (9 points) Dyson Dental Insurance Company sells a dental coverage with the following characteristics:

   i. Each claim is subject to a deductible of 50;
   ii. The maximum amount that can be paid in a year for all claims from a single insured is \$1000.

The number of claims is assumed to follow a binominal distribution with m = 3 and q = 0.7.

The amount of claims is assumed to follow a Weibull distribution with $\theta = 300$ and $\tau = 2$.

The insurance company uses simulation to estimate the claims. The first random number is used to calculate the number of claims. Then the amount of each claim is estimated using the subsequent random numbers using the inverse transformation method.

The random numbers generated from a uniform distribution on (0, 1) are 0.7, 0.1, 0.5, 0.8, 0.3, 0.7, 0.2.

Calculate the simulated amount for this insured that Dyson would have to pay during the year.

# of claims	F(x)
$p_0 = \binom{3}{0}(.7)^0(.3)^{3-0} = .027$	.027
$p_1 = \binom{3}{1}(.7)^1(.3)^2 = .189$	.216
$p_2 = \binom{3}{2}(.7)^2(.3)^1 = .441$	.657
$p_3 = \binom{3}{3}(.7)^3(.3)^0 = .343$	1

First random number is .7 indicating that $N = 3$

Random Number For Amount	X^{**} = Amount	Deduct	Amount PD
0.1	97.38	50	47.38
0.5	249.77	50	199.77
0.8	380.60	50	330.60
			$\boxed{577.74}$

$$u^{**} = F(x^{**}) = 1 - e^{-\left(\frac{x^{**}}{300}\right)^2} \Rightarrow x^{**} = 300\sqrt{-\ln(1-u^{**})}$$

13. (5 points) You are given the following 20 claims:

X: 10, 40, 60, 65, 75, 80, 120, 150, 170, 190, 230, 340, 430, 440, 980, 600, 675, 950, 1250, 1700

The data is being modeled using a Paredo distribution with $\alpha = 4$ and $\theta = 1000$.

Calculate D(350).

$$D(350) = F_n(350) - F^*(350)$$

$$= \frac{12}{20} - \left[1 - \left(\frac{1000}{1350}\right)^4\right]$$

$$= 0.6 - 0.698932$$

$$= -0.098932$$

14. (8 points) A sample of three selected from a distribution produces the following values:

2 2 5

The median of the sample is used to estimate the mean of the distribution.

Estimate the Mean Square Error of this estimate using the Bootstrap Method.

Possibilities	Probability	median $= \hat{\theta}$	$\theta = \bar{X}$
2, 2, 2	$8/27$	2	3
2, 2, 5	$12/27$	2	3
2, 5, 5	$6/27$	5	3
5, 5, 5	$1/27$	5	3

$$MSE = E\left[(\hat{\theta}-\theta)^2\right]$$

$$= \frac{8}{27}(2-3)^2 + \frac{12}{27}(2-3)^2 + \frac{6}{27}(5-3)^2 + \frac{1}{27}(5-3)^2$$

$$= \frac{48}{27} = 1.\overline{77}$$

15. (4 points) Five iPhone's are observed until they fail with failure measured in months. The iPhones have the following failure dates:

8 12 15 25 40

Calculate the variance of $S_5(20)$.

$$S_5(20) = \frac{2}{5}$$

$$Var = \frac{S_5(20)\left[1 - S_5(20)\right]}{n} = \frac{(.4)(.6)}{5} = .048$$

or

$$Var = \frac{n_x(n - n_x)}{n^3} = \frac{(2)(5-2)}{5^3} = \underline{\underline{0.048}}$$