Stat 479
Final
Fall 2009

1. You are given the following data which is assumed to be drawn from a uniform distribution over the range of 0 to ω:

Claim Amount	Count
0 – 200	6
200 – 500	8
500 – 1000	12
1000 – 5000	14
5000 and over	10

Determine maximum likelihood estimate of ω.

$$\hat{W} = \frac{\# \text{ of Observations}}{\# \text{ of observations below } 5000} (5000) = \frac{50}{40}(5000)$$

$$= 6250$$

If you knew that the 10 largest claims ranged from 5,150 to 8000, what would be the maximum likelihood estimate of ω?

$$\hat{W} = \text{max observation} = 8000$$

2. You are given the following sample taken from distribution that follows a Gamma distribution:

$$8 \quad 12 \quad 15 \quad 21$$

Using the Method of Moments, estimate the parameters of the Gamma distribution.

$$E(x) = \alpha \theta = \frac{8 + 12 + 15 + 21}{4} = 14$$

$$E(x^2) = \alpha(\alpha+1)(\theta^2) = \frac{8^2 + 12^2 + 15^2 + 21^2}{4} = 218.50$$

$$\frac{[E(x)]^2}{E(x^2)} = \frac{\alpha^2 \theta^2}{\alpha(\alpha+1)\theta^2} = \frac{\alpha}{\alpha+1} = \frac{(14)^2}{218.50}$$

$$\Rightarrow (14)^2 \alpha + (14)^2 = 218.50\alpha$$

$$\alpha = \frac{196}{22.5} = 8.711$$

$$\theta = \frac{E(x)}{\alpha} = \frac{14}{8.711} = 1.6071$$

3. A sample of claims from Cinfio Casualty Insurance Company is listed below:

500, 600, 800, 1000, 1300, 1800, 2400, 2800, 3000, 5000 $n = 10$
                                    8   9

Frank, the Company's actuary, uses the method of percentile matching at the 80th percentile to determine the parameter θ for an Inverse Pareto distribution assuming that $\tau = 2$.

Determine Frank's estimate of θ.

$$(n+1)(.8) = (11)(.8) = 8.8$$

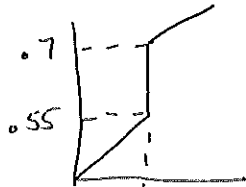$$80^{th} \text{ percentile} = (.2)(2800) + (.8)(3000)$$

$$= 2960$$

$$F(2960) = \left(\frac{2960}{2960 + \theta}\right)^{2} = 0.8$$

$$\theta = 349.3806$$

4. Nick and Vaidehi are covered by a liability insurance policy.

The number of claims for Nick and Vaidehi is assumed to follow a Poison distribution with $\lambda = 2$.

$$\frac{2^k}{k!} e^{-2}$$

The amount of each claim for Nick has the following cdf:

$$F(x) = \begin{cases} 0.001x & 0 \le x < 550 \\ 0.001x + 0.15 & 550 \le x \le 850 \end{cases}$$

N	P_x	$F(x)$
0	$e^{-2} = .1353$	0.1353
1	$2e^{-2} = .27067$	0.4060
2	$2e^{-2} = .27067$	0.6767
3	$\frac{8}{6}e^{-2} = .18044$	0.8571

The amount of each claim for Vaidehi has the following cdf:

$$F(x) = \begin{cases} 0.001x & 0 \le x < 550 \\ 0.55 & 550 \le x < 750 \\ 0.001x - 0.20 & 750 \le x \le 1200 \end{cases}$$

An insurance company uses simulation to estimate the total amount of claims to be incurred by Nick and Vaidehi. First, the company determines the number of claims for Nick and then the amount of each claim for Nick. Next, the company determines the number of claims for Vaidehi. Finally, they simulate the amount of each of Vaidehi 's claims.

The random numbers generated for the simulation are:

$N=2$ | $N=3$
0.50 0.30 0.65 | 0.75 0.70 0.55 0.25 0.60 0.45 0.15

Determine the total amount claims paid for each of Nick and Vaidehi.

NICK

Random Number

.5 ⟹ 2 claims

Amount

.3 ⟹ $0.001x_1 = .3$

$x_1 = 300$

.65 ⟹ $x_2 = 550$

Total $= x_1 + x_2 = 850$

Vaidehi

Random number

.75 ⟹ 3 claims

Amount

.7 $.001x_1 - 0.2 = .7$

$x_1 = 900$

.55 $x_2 = 750$ Use largest value

.25 $.001x_3 = .25$

$x_3 = 250$

Total $= 900 + 750 + 250 = 1900$

5. To estimate E(X), you have simulated five observations from the random variable X.

The values are 1, 3, 5, 8, and 8.

Your goal is to have the standard deviation of the estimate of E(X) be less than 0.5.

Estimate the total number of simulations needed.

$$E(X) = \frac{1+3+5+8+8}{5} = 5$$

$$S^2 = \frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (8-5)^2 + (8-5)^2}{4}$$

$$= 9.5$$

$$SD(\bar{X}) = \sqrt{\frac{S^2}{n}} < 0.5$$

$$= \sqrt{\frac{9.5}{n}} < 0.5$$

$$\Rightarrow n > 38$$

6. The mean of a distribution will be estimated by taking the average of the smallest value and the largest value from a sample. In other words,

$$\hat{\mu} = 0.5[\text{Min}(X_1, X_2, ..., X_n) + \text{Max}(X_1, X_2, ..., X_n)]$$

The following sample is selected from the distribution:

4, 4, 7

Using the bootstrap method, estimate the mean square error in the above estimator.

Possible Draws	Prob	$\hat{\mu}$	TRUE μ
4, 4, 4	$(\frac{2}{3})(\frac{2}{3})(\frac{2}{3}) = \frac{8}{27}$	.4	5
4, 4, 7	$(\frac{2}{3})(\frac{2}{3})(\frac{1}{3})(3) = \frac{12}{27}$	5.5	5
4, 7, 7	$(\frac{2}{3})(\frac{1}{3})(\frac{1}{3})(3) = \frac{6}{27}$	5.5	5
7, 7, 7	$\frac{1}{27}$	7	5

$$MSE(\hat{\mu}) = \frac{8}{27}(5-4)^2 + \frac{12}{27}(5-5.5)^2$$
$$+ \frac{6}{27}(5-5.5)^2 + \frac{1}{27}(5-7)^2$$

$$= \frac{8}{27} + \frac{3}{27} + \frac{1.5}{27} + \frac{4}{27}$$

$$= \frac{16.5}{27} = 0.6111\overline{1}$$

7. An insurance policy pays claims up to a limit of 2000. A random sample of two payments is obtained as follows: 300 and 2000.

The claims are assumed to follow an Inverse Pareto distribution with $\tau = 1$.

Calculate the maximum likelihood estimate for θ.

$$L(\theta) = f(300)\left[1 - F(2000)\right]$$

$$= \frac{\theta}{(300+\theta)^2}\left[1 - \frac{2000}{2000+\theta}\right] = \frac{\theta}{(300+\theta)^2} \cdot \frac{\theta}{2000+\theta}$$

$$\ell(\theta) = 2\ln\theta - 2\ln(300+\theta) - \ln(2000+\theta)$$

$$\ell'(\theta) = \frac{2}{\theta} - \frac{2}{300+\theta} - \frac{1}{2000+\theta} = 0$$

$$\frac{2(300+\theta)(2000+\theta) - 2(\theta)(2000+\theta) - \theta(300+\theta)}{(\theta)(300+\theta)(2000+\theta)} = 0$$

$$(600+2\theta)(2000+\theta) - 4000\theta - 2\theta^2 - 300\theta - \theta^2 = 0$$

$$1,200,000 + 4600\theta + 2\theta^2 - 4300\theta - 3\theta^2 = 0$$

$$\theta^2 - 300\theta - 1,200,000 = 0$$

$$\theta = \frac{300 \pm \sqrt{(300)^2 - (4)(1)(-1,200,000)}}{2}$$

$$= 1255.6672$$

8. Claims for a hospital stay are assumed to follow an exponential distribution.

Hong Insurance offers a hospitalization policy that covers hospital claims up to a maximum of 10,000. A sample of ten claims is drawn to estimate θ. The ten claims in the sample are:

500 700 850 1200 1500 2500 5000 7800 10,000 10,000

Estimate θ using the maximum likelihood estimate.

$$\theta = \frac{\Sigma \text{ of observations}}{\text{\# of uncensored}} =$$

$$\frac{500 + 700 + 850 + 1200 + 1500 + 2500 + 5000 + 7800 + 20000}{8}$$

$$= 5006.25$$

9. You have a sample of 19 claims for dental insurance claims. Using the smoothed empirical distribution, the 42nd percentile is 207.4 and the 44th percentile is 215.8.

Determine X_8 and X_9.

$$n = 19$$

$$(n+1)(.42) = 8 \cdot 4 \quad \Rightarrow \quad .6 X_8 + .4 X_9 = 207.40 \quad ①$$

$$(n+1)(.44) = 8.8 \quad \Rightarrow \quad .2 X_8 + .8 X_9 = 215.8 \quad ②$$

$$1.2 X_8 + .8 X_9 = 414.8 \quad ① \times 2$$

$$X_8 = 414.8 - 215.8 \quad [① \times 2] - ②$$

$$X_8 = 199$$

$$X_9 = 220$$

10. During a one-year period, Ren Wreckers suffered the following liability claims:

400 900 1400 3300

The losses are assumed to be distributed following a Pareto distribution with $\alpha = 4$ and $\theta = 5000$.

Xuelu wants to test this hypothesis using the Kolmogorov-Smirnov test with a significance level of 5%.

Calculate the Kolmogorov-Smirnov test statistic.

X	$F(x-)$	$F(x)$	$F^*(x) = 1 - \left(\dfrac{5000}{x+5000}\right)^4$	$MAX\left(F(x) - F^*(x)\right)$
400	0	.25	0.26497	0.26497
900	.25	.50	0.48421	0.23421
1400	.50	.75	0.62747	0.12747
3300	.75	1.00	0.86831	0.13169

$$D = 0.26497$$

State the critical value for the test.

At 5% Significance Level

Critical Value $= \dfrac{1.36}{\sqrt{n}} = \dfrac{1.36}{\sqrt{4}} = 0.68$

State your conclusion.

$0.26 < 0.68$

Do Not Reject H_0

11. During a one-year period, Ren Wreckers suffered the following liability claims:

$$400 \quad 900 \quad 1400 \quad 3300$$

The losses are assumed to be distributed following an Pareto distribution with $\alpha = 4$ and $\theta = 5000$.

You want to create a *p-p* plot to evaluate the assumption.

Determine the plotted point for the observed value of 900.

$$point = \left[\frac{i}{n+1}, F^*(x_i) \right]$$

$$F^*(900) = 1 - \left(\frac{5000}{5900} \right)^4 = .4842$$

$$so \quad point = \left[\frac{2}{5}, .4842 \right]$$

12. During an hour, the number of cars that arrive at a gas station during any one minute is given below:

Number of Cars Arriving	Count
0	8
1	12
2	15
3	10
4	7
5	6
6	0
7	1
8	1

$$\Sigma = 60$$

Assuming that arrivals are distributed as a Poisson distribution, determine the 90% linear confidence interval for λ, the expected number of cars arriving at the gas station in any one minute period.

$$E(\lambda) = \bar{X} = \frac{12 + (15)(2) + (10)(3) + (7)(4) + 5(6) + 7 + 8}{60}$$

$$= 2.4167$$

$$Var(\lambda) = \frac{E(\lambda)}{n} = \frac{2.4167}{60} = .0403$$

$$90\% \ C.I. = E(\lambda) \pm 1.645 \sqrt{Var(\lambda)}$$

$$= 2.4167 \pm 1.645 \sqrt{.0403}$$

$$= [2.0865, 2.7468]$$

13. During the last year, Edson Fire Insurance Company incurred 20 claims. The claims were distributed as follows:

Amount of Claim	Count
0 – 10,000	10
10,000 – 25,000	7
25,000 – 100,000	2
100,000 +	1

$\sum = 20$

The Company wants to test their belief that the claims are distributed exponentially with a mean of 20,000. Therefore,

H_0: Fire Claims are distributed exponentially with a mean of 20,000.

H_1: Fire Claims are not distrusted exponentially with a mean of 20,000.

Calculate the Chi Square test statistic.

$$E_j = \left(e^{-\frac{x_{i-1}}{\theta}} - e^{-\frac{x_i}{\theta}} \right) n$$

Range	O_j		$(E_j - O_j)^2 / E_j$
0 – 10,000	10	$\left(1 - e^{-\frac{10}{20}}\right)20 = 7.8694$	0.5768
10,000 – 25,000	7	$\left(e^{-\frac{10}{20}} - e^{-\frac{25}{20}}\right)(20) = 6.4005$	0.0561
25,000 – 100,000	2	$\left(e^{-\frac{25}{20}} - e^{-\frac{100}{20}}\right)(20) = 5.5953$	2.3102
100,000 +	1	$\left(e^{-\frac{100}{20}}\right)(20) = 0.13476$	5.5554

$$\chi^2 = 8.499$$

Calculate the critical values for the Chi Square test at a 5% significance level.

degrees of freedom = 4 - 1 - 0 = 3

critical value = 7.815

State your conclusion with regard to Edson's hypothesis.

$$\chi^2 = 8.5 > 7.815$$

Reject H_0

14. A golf course buys an insurance policy from Amstutz Assurance Company which covers damage to adjacent houses caused by errant golf balls. The policy covers each claim after a deductible of 50 for each claim.

The number of claims during a round of golf follows a binomial distribution with m = 3 and q = 0.6

The amount of each claim follows a Pareto distribution with α = 11 and θ = 1000. $F(x) = 1 - \left(\frac{1000}{1000+x}\right)^{11}$

Using simulation, Alex wants to estimate the total claims that will need to be paid by Amstutz Assurance Company. He does so by estimating the claims for each golfer. John and Ian are the first two golfers. First, Alex determines the number of claims for John and then the amount of each claim for John. Next, Alex determines the number of claims for Ian. Finally, Alex simulates the amount of each of Ian's claims.

The random numbers used in the simulation are:

JOHN CLAIM ↓

0.10 0.85 | 0.50 0.30 0.90 0.45 0.60 0.95 0.75 0.05

↑ NUMBER OF JOHN'S CLAIMS

Calculate the amount the Amstutz will have to pay for as a result of Ian's round of golf.

DISTRIBUTION OF # of CLAIMS

$$\binom{3}{k}(0.6)^k (0.4)^{3-k}$$

→

#	$f(x)$	$F(x)$
0	0.064	0.064
1	0.288	0.352
2	0.432	0.784
3	0.216	1

JOHN
# of claims   0.1 ⟹ 1 claim
Amount ⟹ NOT IMPORTANT SINCE WE WANT IAN'S COST
↓

IAN
# of Claims   0.5 ⟹ 2 claims

AMOUNT

RANDOM NUMBER
0.3 $1 - \left(\frac{1000}{1000+X_1}\right)^{11} = .3$

$X_1 = 32.9564$

0.9 $1 - \left(\frac{1000}{1000+X_2}\right) = .9$

$X_2 = 232.8467$

APPLY $50 Deductible

Claim 1 = MAX(0, X_1 - 50) = 0
CLAIM 2 = MAX(0, X_2 - 50) = 182.8467

TOTAL = 0 + 182.8467
      = 182.8467

15. You are given the following sample taken from distribution that follows a Binomial distribution:

$$6 \quad 7 \quad 9 \quad 11 \quad 12$$

Using the Method of Moments, estimate the parameters of the Binomial distribution.

$$\bar{X} = \frac{6+7+9+11+12}{5} = 9$$

$$E(X^2) = \frac{36+49+81+121+144}{5} = 86.2$$

$$Var = 86.2 - (9)^2 = 5.2$$

$$m\,g = 9$$

$$m\,g\,(1-g) = 5.2$$

$$g = 0.4222$$

$$m = 21.3$$

Round m to nearest integer

$$\hat{m} = 21$$

$$\hat{g} = \frac{9}{21} = .42857$$