

Name:

M339J/M389J: Probability Models with Actuarial Applications

The University of Texas at Austin

Sample Problems for In-Term Exam III

Instructor: Milica Čudina

**Notes:** This is a closed book and closed notes exam. The maximal score on this exam is 100 points.

---

**3.1. Free-response problems.** *Please, explain carefully all your statements and assumptions. Numerical results or single-word answers without an explanation (even if they're correct) are worth 0 points.*

**Problem 3.1.** A Poisson distribution is used to fit frequency data. You are given that:

- there was a total of 20 observations less than or equal to 1;
- there were 6 observations equal to 2;
- there were 4 observations equal to 3;
- there were no observations greater than or equal to 4.

What is the maximum likelihood estimate for the mean of this Poisson distribution?

**Solution:** The likelihood function is

$$L(\lambda) = (e^{-\lambda}(1 + \lambda))^{20} (e^{-\lambda} \frac{\lambda^2}{2!})^6 (e^{-\lambda} \frac{\lambda^3}{3!})^4 \propto e^{-30\lambda} (1 + \lambda)^{20} \lambda^{24}.$$

The likelihood function is

$$l(\lambda) = \text{const} - 30\lambda + 20 \ln(1 + \lambda) + 24 \ln(\lambda).$$

Differentiating, we obtain

$$l'(\lambda) = -30 + \frac{20}{1 + \lambda} + \frac{24}{\lambda}.$$

We equate the above to zero and solve for  $\lambda$  in

$$-15\lambda(1 + \lambda) + 10\lambda + 12(1 + \lambda) = 0,$$

i.e.,

$$15\lambda^2 - 7\lambda - 12 = 0 \quad \Rightarrow \quad \lambda = \frac{7 + \sqrt{49 + 720}}{30} = 1.157695.$$

**Problem 3.2.** (10 points) In the compound model for aggregate claims, let the frequency random variable  $N$  have the Poisson distribution with mean 1.

Let the common distribution of the i.i.d. severity random variables  $\{X_j; j = 1, 2, \dots\}$  be given by the following p.m.f.

$$p_X(100) = 1/2, p_X(200) = 3/10, p_X(300) = 1/5.$$

Let our usual assumptions hold, i.e., let  $N$  be independent of  $\{X_j; j = 1, 2, \dots\}$ .

Define the aggregate loss as  $S = \sum_{j=1}^N X_j$ .

Find the expected value of the **policyholder's** payment for a stop-loss insurance policy with an ordinary deductible of 200, i.e., calculate  $\mathbb{E}[S \wedge 200]$ .

**Solution:** Note that  $S$  has the support of the form  $\{0, 100, 200, 300, \dots\}$ . So,

$$\mathbb{E}[S \wedge 200] = 100\mathbb{P}[S = 100] + 200\mathbb{P}[S \geq 200].$$

Next,

$$\mathbb{P}[S = 0] = \mathbb{P}[N = 0] = e^{-1},$$

$$\mathbb{P}[S = 100] = \mathbb{P}[N = 1, X_1 = 100] = 0.5e^{-1},$$

$$\mathbb{P}[S \geq 200] = 1 - \mathbb{P}[S = 0] - \mathbb{P}[S = 100] = 1 - 1.5e^{-1}.$$

So,

$$\mathbb{E}[S \wedge 200] = 100 \cdot 0.5e^{-1} + 200(1 - 1.5e^{-1}) = 200 - 250e^{-1} \approx 108.03.$$

**Problem 3.3.** (10 points) The frequency random variable  $N$  is assumed to have a Poisson distribution with a mean of 2. Individual claim severity random variable  $X$  has the following probability mass function

$$p_X(100) = 0.6, \quad p_X(200) = 0.3, \quad p_X(300) = 0.1.$$

Let the above be the common distribution of the i.i.d. severity random variables  $\{X_j; j = 1, 2, \dots\}$ , and Let our usual assumptions hold, i.e., let  $N$  be independent of  $\{X_j; j = 1, 2, \dots\}$ . Define the aggregate loss as  $S = \sum_{j=1}^N X_j$ . Calculate the probability that  $S$  is exactly equal to 3.

**Solution:**

$$\begin{aligned} \mathbb{P}[S = 300] &= \mathbb{P}[N = 1, X_1 = 300] + \mathbb{P}[N = 2, X_1 = 100, X_2 = 200] \\ &\quad + \mathbb{P}[N = 2, X_1 = 200, X_2 = 100] + \mathbb{P}[N = 3, X_1 = 100, X_2 = 100, X_3 = 100] \\ &= 2e^{-2} \cdot 0.1 + 2 \cdot \frac{2^2}{2} e^{-2} \cdot 0.6 \cdot 0.3 + \frac{2^3}{3 \cdot 2} \cdot e^{-2} \cdot 0.6^3 \\ &= e^{-2}(0.2 + 0.72 + 0.288) = 1.208e^{-2} \approx 0.1635. \end{aligned}$$

**Problem 3.4.** (6 points) In the compound model for aggregate claims, let the frequency random variable  $N$  have the probability (mass) function

$$p_N(0) = 0.4, p_N(1) = 0.3, p_N(2) = 0.2, p_N(4) = 0.1.$$

Moreover, let the common distribution of the i.i.d. severity random variables  $\{X_j; j = 1, 2, \dots\}$  be given by the probability (mass) function  $p_X(1) = 0.3$  and  $p_X(2) = 0.7$ .

Let our usual assumptions hold, i.e., let  $N$  be independent of  $\{X_j; j = 1, 2, \dots\}$ .

Define the aggregate loss as  $S = \sum_{j=1}^N X_j$ .

Calculate  $\mathbb{E}[(S - 2)_+]$ .

**Solution:** We intend to use the equality

$$\mathbb{E}[(S - 2)_+] = \mathbb{E}[S] - \mathbb{E}[S \wedge 2].$$

Using

$$\begin{aligned}\mathbb{E}[N] &= 0.4 \cdot 0 + 0.3 \cdot 1 + 0.2 \cdot 2 + 0.1 \cdot 4 = 0.3 + 0.4 + 0.4 = 1.1, \\ \mathbb{E}[X] &= 0.3 \cdot 1 + 0.7 \cdot 2 = 0.3 + 1.4 = 1.7,\end{aligned}$$

we get

$$\mathbb{E}[S] = \mathbb{E}[N]\mathbb{E}[X] = 1.1 \cdot 1.7 = 1.87.$$

On the other hand,

$$\begin{aligned}\mathbb{E}[S \wedge 2] &= \mathbb{P}[S > 0] + \mathbb{P}[S > 1] \\ &= (1 - F_S(0)) + (1 - F_S(1)).\end{aligned}$$

From the problem statement, we conclude that

$$\begin{aligned}F_S(0) &= \mathbb{P}[S \leq 0] = \mathbb{P}[S = 0] = \mathbb{P}[N = 0] = 0.4, \\ F_S(1) &= \mathbb{P}[S \leq 1] = \mathbb{P}[S = 0] + \mathbb{P}[S = 1] = \mathbb{P}[N = 0] + \mathbb{P}[N = 1, X_1 = 1] \\ &= 0.4 + 0.3 \cdot 0.3 = 0.49.\end{aligned}$$

Finally,

$$\mathbb{E}[S \wedge 2] = 0.6 + 0.51 = 1.11.$$

and

$$\mathbb{E}[(S - 2)_+] = 1.87 - 1.11 = 0.76.$$

**Problem 3.5.** (15 points) A sizeable village of log cabins is insured. Aggregate losses have a compound Poisson distribution. The expected number of losses is 25 per year. Loss amounts, regardless of the type of wood used to construct the cabin, follow an exponential distribution with mean 2000.

The insurer seeks to reduce the cost of insurance so they decide to do two things:

- They stop insuring pine cabins which make up one fifth of the total number of cabins.
- They impose an ordinary deductible of 500 per loss.

What are the expected value and the variance of the total aggregate amount paid by the insurer after the above modifications are implemented?

**Solution:** The old loss frequency is  $N_L \sim \text{Poisson}(\lambda_L = 25)$ . So, the new loss frequency can be modeled as  $\tilde{N}_L \sim \text{Poisson}(\tilde{\lambda}_L = 25(4/5) = 20)$ .

The loss distribution is given to be  $X \sim \text{Exponential}(\theta = 2000)$ . After the ordinary deductible is imposed, the per loss random variable has the following distribution

$$Y^L \sim \begin{cases} 0 & \text{with probability } F_X(500) \\ Y^P & \text{with probability } S_X(500) \end{cases}$$

with  $Y^P \sim \text{Exponential}(\theta = 2000)$ .

The aggregate claims  $S = Y_1^L + Y_2^L + \cdots + Y_{\tilde{N}_L}^L$  have the mean

$$\mathbb{E}[S] = \mathbb{E}[Y^L]\mathbb{E}[\tilde{N}_L] = 2000e^{-500/2000}(20) = 31152.03.$$

The variance of the claims is

$$\text{Var}[S] = \mathbb{E}[\tilde{N}_L]\text{Var}[Y^L] + \text{Var}[\tilde{N}_L](\mathbb{E}[Y^L])^2 = \tilde{\lambda}_L\mathbb{E}[(Y^L)^2] = (20)(2(2000)^2)e^{-500/2000} = 124608125.$$

### 3.2. MULTIPLE CHOICE QUESTIONS.

**Problem 3.6.** (5 points) Let  $S$  be the aggregate claims random variable. You are given the following:

- (i)  $\mathbb{E}[(S - 100)_+] = 15$ ,
- (ii)  $\mathbb{E}[(S - 120)_+] = 10$ ,
- (iii)  $\mathbb{P}[80 < S \leq 120] = 0$ .

Find  $\mathbb{E}[(S - 105)_+]$ .

- (a)  $\mathbb{E}[(S - 105)_+] \approx 10.75$
- (b)  $\mathbb{E}[(S - 105)_+] \approx 12.75$
- (c)  $\mathbb{E}[(S - 105)_+] \approx 13.75$
- (d)  $\mathbb{E}[(S - 105)_+] \approx 15.75$
- (e) None of the above

**Solution: (c)**

By the Interpolation Theorem,

$$\mathbb{E}[(S - 105)_+] = \frac{3}{4}\mathbb{E}[(S - 100)_+] + \frac{1}{4}\mathbb{E}[(S - 120)_+] = 11.25 + 2.5 = 13.75.$$

**Problem 3.7.** (5 points) Consider the following individual observed values:

$$5, 8, 10$$

of a random variable  $X$  whose distribution function is given by  $F_X(x) = 1 - (1/x)^p$  for  $x > 1$  and an unknown parameter  $p$ . Let  $\hat{p}$  denote the Maximum Likelihood Estimate of  $p$  based on the above observed values. Then,

- (a)  $\hat{p} \approx \frac{3}{2}$
- (b)  $\hat{p} \approx \frac{3 \ln(20)}{2}$
- (c)  $\hat{p} \approx \frac{3 \ln(20)}{4}$
- (d)  $\hat{p} \approx \frac{3}{2 \ln(20)}$
- (e) None of the above

**Solution: (d)**

The density function of  $X$  is

$$f_X(x) = F'_X(x) = -(-p)x^{-p-1} = px^{-p-1} \quad x > 1.$$

So, the likelihood function has the form

$$\begin{aligned} L(p) &= p5^{-p-1} \cdot p8^{-p-1} \cdot p10^{-p-1} \\ &= p^3(5 \cdot 8 \cdot 10)^{-p-1}, \end{aligned}$$

and the loglikelihood function is

$$l(p) = 3\ln(p) - (p+1)\ln(400).$$

Differentiating with respect to  $p$ , we get

$$l'(p) = 3/p - \ln(400).$$

Setting the above equal to zero and solving for  $p$ , we get

$$\hat{p} = 3/\ln(400) = 3/(2\ln(20)).$$

**Problem 3.8.** (5 points) Consider the following individual observed values

$$5, 10$$

and the one right censored value 8 of a random variable  $X$  whose distribution function is given by  $F_X(x) = 1 - (1/x)^p$  for  $x > 1$  and an unknown parameter  $p$ . Let  $\hat{p}$  denote the Maximum Likelihood Estimate of  $p$  based on the above observed values. Then,

- (a)  $\hat{p} \approx \frac{1}{\ln(20)}$
- (b)  $\hat{p} \approx \frac{3\ln(20)}{2}$
- (c)  $\hat{p} \approx \frac{3}{4\ln(20)}$
- (d)  $\hat{p} \approx \frac{3}{2\ln(20)}$
- (e) None of the above

**Solution: (a)**

The density function of  $X$  is

$$f_X(x) = F'_X(x) = -(-p)x^{-p-1} = px^{-p-1} \quad x > 1.$$

So, the likelihood function has the form

$$\begin{aligned} L(p) &= p5^{-p-1} \cdot 8^{-p} \cdot p10^{-p-1} \\ &= p^2(5 \cdot 10)^{-p-1}8^{-p}, \end{aligned}$$

and the loglikelihood function is

$$l(p) = 2\ln(p) - (p+1)\ln(50) - p\ln(8).$$

Differentiating with respect to  $p$ , we get

$$l'(p) = 2/p - \ln(400).$$

Setting the above equal to zero and solving for  $p$ , we get

$$\hat{p} = 2/\ln(400) = 2/(2\ln(20)) = 1/\ln(20).$$

**Problem 3.9.** You fit a distribution with the following density function:

$$f_X(x) = (p+1)x^p, \quad 0 < x < 1, p > -1.$$

As usual, your observations are denoted by

$$x_1, x_2, \dots, x_n.$$

What is the expression for the maximum likelihood estimate of the parameter  $p$ ?

- (a)  $-\frac{n}{\sum_{i=1}^n \ln(x_i)} - 1$
- (b)  $-\frac{1}{n} \sum_{i=1}^n \ln(x_i)$
- (c)  $\frac{1}{n} \sum_{i=1}^n x_i$
- (d)  $\frac{\sum_{i=1}^n \ln(x_i)}{n} - 1$
- (e) None of the above.

**Solution: (a)**

The likelihood function is

$$L(p) = \prod_{i=1}^n f_X(x_i; p) = \prod_{i=1}^n ((p+1)x_i^p) = (p+1)^n \left( \prod_{i=1}^n x_i \right)^p.$$

Hence, the log-likelihood function is

$$\ell(p) = \ln(L(p)) = n \ln(p+1) + p \left( \sum_{i=1}^n \ln(x_i) \right).$$

The derivative of the log-likelihood function is

$$\ell'(p) = \frac{n}{p+1} + \left( \sum_{i=1}^n \ln(x_i) \right).$$

Equating the above to zero, and solving for  $p$ , we obtain

$$\frac{n}{p+1} = - \sum_{i=1}^n \ln(x_i) \quad \Rightarrow \quad \hat{p}_{MLE} = - \frac{n}{\sum_{i=1}^n \ln(x_i)} - 1.$$

**Problem 3.10.** You have observed the following three loss amounts:

$$190, \quad 90, \quad 60$$

Four other loss amounts are known to be less than or equal to 60. Losses follow an inverse exponential distribution. Calculate the maximum likelihood estimate of the population mode based on the above data.

- (a) 10.125
- (b) 12.378
- (c) 15.044
- (d) 20.232
- (e) None of the above.

**Solution: (c)**

The cumulative distribution function of the inverse exponential distribution is given in the STAM tables:

$$F_X(x; \theta) = e^{-\theta/x},$$

while its probability density function has the form:

$$f_X(x; \theta) = \frac{\theta e^{-\theta/x}}{x^2}, \quad x > 0,$$

So, the likelihood function in this problem has the form

$$\begin{aligned} L(\theta) &= f_X(190; \theta) f_X(90; \theta) f_X(60; \theta) (F_X(x; \theta))^4 \\ &\propto (\theta e^{-\theta/190})(\theta e^{-\theta/90})(\theta e^{-\theta/60}) (e^{-\theta/60})^4 \\ &= \theta^3 e^{-\theta(\frac{1}{190} + \frac{1}{90} + \frac{1}{60} + \frac{1}{15})}. \end{aligned}$$

Thus, the log-likelihood function can be written as

$$\ell(\theta) = c + 3 \ln(\theta) - \theta \left( \frac{1}{190} + \frac{1}{90} + \frac{1}{60} + \frac{1}{15} \right)$$

where  $c$  stands for an irrelevant constant. Differentiating the log-likelihood function with respect to  $\theta$ , we obtain

$$\ell'(\theta) = \frac{3}{\theta} - \left( \frac{1}{190} + \frac{1}{90} + \frac{1}{60} + \frac{1}{15} \right).$$

We equate the above expression to zero and solve for  $\theta$ . We get

$$\frac{3}{\theta} = \frac{1}{190} + \frac{1}{90} + \frac{1}{60} + \frac{1}{15} \quad \Rightarrow \quad \hat{\theta}_{MLE} = \frac{3}{\frac{1}{190} + \frac{1}{90} + \frac{1}{60} + \frac{1}{15}} = 30.08797654.$$

According to the STAM tables, the mode of the inverse exponential distribution is  $\theta/2$ . So, our answer is 15.044.

**Problem 3.11.** *Source: Sample STAM Problem #196.*

You are given the following 10 bodily injury losses (before the deductible is applied):

Loss amount	Number of losses	Policy limit
200	3	500
400	4	800
> 800	3	800

Past experience indicates that these losses follow a two-parameter Pareto distribution with parameters  $\alpha$  unknown and  $\theta = 1,000$ . Calculate the maximum likelihood estimate of  $\alpha$ .

- (a) 1.9145
- (b) 2.307
- (c) 2.853
- (d) 3.089
- (e) None of the above.

**Solution: (a)**

According to the STAM tables, the density function and the distribution function of the two-parameter Pareto are

$$f_X(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}} \quad \text{and} \quad F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha.$$

So, the likelihood function with the above 7 unmodified data points and 3 right-censored data points is

$$\begin{aligned} L(\alpha) &= \left(\frac{\alpha 1000^\alpha}{(200+1000)^{\alpha+1}}\right)^3 \left(\frac{\alpha 1000^\alpha}{(400+1000)^{\alpha+1}}\right)^4 \left(\frac{1000}{800+1000}\right)^{3\alpha} \\ &= \left(\frac{\alpha 1000^\alpha}{(1200)^{\alpha+1}}\right)^3 \left(\frac{\alpha 1000^\alpha}{(1400)^{\alpha+1}}\right)^4 \left(\frac{1000}{1800}\right)^{3\alpha} \\ &\propto \alpha^7 (1000)^{10\alpha} (1200)^{-3\alpha} (1400)^{-4\alpha} (1800)^{-3\alpha}. \end{aligned}$$

So, the log-likelihood can be written as

$$\ell(\alpha) = 7 \ln(\alpha) + 10\alpha \ln(1000) - 3\alpha(\ln(1200) + \ln(1800)) - 4\alpha \ln(1400) + c$$

where  $c$  denotes an irrelevant constant. Differentiating with respect to  $\alpha$ , we obtain

$$\ell'(\alpha) = \frac{7}{\alpha} + 10 \ln(1000) - 3(\ln(1200) + \ln(1800)) - 4 \ln(1400).$$

Equating the above expression to zero, and solving for  $\alpha$ , we get

$$\alpha = \frac{7}{-10 \ln(1000) + 3(\ln(1200) + \ln(1800)) + 4 \ln(1400)} = 1.914548969.$$

**Problem 3.12.** *Source: Sample STAM Problem #152.*

You are given the following information:



- (i) A sample of losses is: 400, 600, 700, 900, 1000
- (ii) No information is available about losses of 300 or less.
- (iii) Losses are assumed to follow an exponential distribution with mean  $\theta$ .

Calculate the maximum likelihood estimate of  $\theta$  based on the above data.

- (a) 420
- (b) 520
- (c) 720
- (d) 920
- (e) None of the above.

**Solution: (a)**

This is a case of truncated data. We have worked out in class that in the exponential case with data truncated at  $d$ , the expression for the maximum likelihood estimate of  $\theta$  reads as

$$\hat{\theta}_{MLE} = \bar{x} - d.$$

With the given data, we get

$$\hat{\theta}_{MLE} = \frac{400 + 600 + 700 + 900 + 1000}{5} - 300 = 420.$$