HOMEWORK 5

Problem 1: (Best Unbiased estimators, 20pts)

1. Let $X_1, \dots, X_n$ be iid Poisson Random variables with parameter $\lambda > 0$. Find the UMVUE for (i) $e^{-\lambda}$ (ii) $\lambda e^{-\lambda}$.

Solution: Poisson r.v. $X_i \sim \frac{\lambda^k e^{-\lambda}}{k!}$

We know that $T = \sum_{i=1}^{n} X_i$ is a complete sufficient statistic for λ.

Also T is poisson$(n\lambda)$

By Lehmann-Scheffé theorem, UMVUE for any parameter $\theta(\lambda)$ can be obtained by first finding an unbiased estimator U of $\theta(\lambda)$ and then conditioning that estimator on T to get $\hat{\theta} = \mathbb{E}[U|T]$ which is UMVUE.

(i) We need to find a statistic U for $e^{-\lambda}$.

Since X_i are iid & $X_i \sim \frac{\lambda^k e^{-\lambda}}{k!}$, if we take $k=0$,

we get $P(X_i = 0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda}$.

So we can pick any particular $i \in [0, n]$ & $U = \mathbb{I}[X_i = 0]$.

Let us pick $i=1$ for simplicity. Then $U = \mathbb{I}[X_i = 0]$.

U is an unbiased estimator of $e^{-\lambda}$ because $E[U] = P(X_i = 0) = e^{-\lambda}$.

To get UMVUE, we take the conditional expectation $\mathbb{E}[U|T]$ where T is a complete sufficient statistic. We know $T = \sum_{i=1}^{n} X_i$ is complete.

$$E[U|T=t] = \sum_u u\, p(u|T=t) = 0 \cdot p(u=0 \mid T=t) + 1 \cdot P(u=1 \mid T=t)$$
$$= P(u=1 \mid T=t) = P(X_1 = 0 \mid T=t)$$

Now $P(X_1=0 \mid T=t) = P(X_1=0 \mid X_1+X_2+\cdots X_n=t)$

$$= \frac{P(X_1=0,\ X_1+\cdots X_n=t)}{P(X_1+X_2+\cdots X_n=t)}$$

$$= \frac{P(X_1=0,\ X_2+X_3+\cdots X_n=t)}{P(X_1+X_2+\cdots X_n=t)}$$

$$= \frac{P(X_1=0)\ P(X_2+X_3+\cdots X_n=t)}{P(X_1+\cdots X_n=t)}$$

$$= \frac{\dfrac{e^{-\lambda}\lambda^0}{0!} \cdot \dfrac{e^{-(n-1)\lambda}[(n-1)\lambda]^t}{\cancel{t!}}}{\dfrac{e^{-n\lambda}[n\lambda]^t}{\cancel{t!}}}$$

$$= \frac{e^{-\lambda}\cdot e^{-(n-1)\lambda}\cdot (n-1)^t \lambda^t}{e^{-n\lambda}(n\lambda)^t}$$

$$= \frac{\cancel{e^{-n\lambda}}\ (n-1)^t\ \cancel{\lambda^t}}{\cancel{e^{-n\lambda}}\ n^t\ \cancel{\lambda^t}}$$

$$= \left(\frac{n-1}{n}\right)^t = \left(1-\frac{1}{n}\right)^t$$

$\therefore\ E[U \mid T=t] = \left(\frac{n-1}{n}\right)^t$

or $E[U \mid T] = \left(\frac{n-1}{n}\right)^T = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^n x_i}$

$\therefore$ the UMVUE is $\widehat{e^{-\lambda}} = E[U \mid T] = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^n x_i}$

(ii) Now we want to find the UMVUE for (ii) $\lambda e^{-\lambda}$.

We can see, just like in (i), $P(X_1 = 1) = \frac{e^{-\lambda}\lambda^1}{1!} = \lambda e^{-\lambda}$

So we can just take $U = \mathbb{I}_{[X_1 = 1]}$ such that

$$E[U] = P(X_1 = 1) = \lambda e^{-\lambda}$$

which makes U an unbiased estimator of $\lambda e^{-\lambda}$.

We can do the same thing we did in the previous part.

$$E[U|T=t] = \sum_u u\, p(u=u|T=t)$$

$$= 1 \cdot p(u=1|T=t) + 0 \cdot P(u=0|T=t)$$

$$= p(u=1|T=t)$$

The event $U=1$ is same as the event $X_1 = 1$.

$$\therefore E[U|T=t] = p(u=1|T=t) = P(X_1 = 1|T=t)$$

$$= \frac{P(X_1=1, T=t)}{P(T=t)} = \frac{P(X_1=1, X_1+X_2+\cdots X_n = t)}{P(T=t)}$$

$$= \frac{P(X_1=1, X_2+X_3+\cdots X_n = t-1)}{P(X_1+X_2+\cdots+X_n = t)}$$

$$= \frac{P(X_1=1)\, P(X_2+X_3+\cdots X_n = t-1)}{P(X_1+X_2+\cdots X_n = t)}$$

$$\frac{\frac{\lambda^1 e^{-\lambda}}{1!} \cdot \frac{e^{-(n-1)\lambda} \cdot [(n-1)\lambda]^{t-1}}{(t-1)!}}{\frac{e^{-n\lambda}(n\lambda)^t}{t!}} = \frac{\lambda \cancel{e^{-\lambda}} \cancel{e^{-(n-1)\lambda}} \cdot (n-1)^{t-1} \lambda^{t-1}}{\frac{\cancel{e^{-n\lambda}}\, n^t \lambda^t}{t!} \cdot (t-1)!}$$

$$= \frac{\lambda \cdot (n-1)^{t-1} \lambda^{t-1}}{\frac{n^t \lambda^t}{t!} \quad (t-1)!} = \frac{\cancel{\lambda} \cdot (n-1)^{t-1} \cancel{\lambda^{t-1}}}{\frac{n^t \cancel{\lambda^t}}{t}}$$

$$= \frac{t}{n} \left(\frac{n-1}{n}\right)^{t-1} = \frac{t}{n-1} \left(\frac{n-1}{n}\right)^t$$

$$\therefore \widehat{\lambda e^{-\lambda}} = \frac{T}{n}\left(\frac{n-1}{n}\right)^{T-1} = \frac{T}{n-1}\left(\frac{n-1}{n}\right)^T$$

where $T = \sum_{i=1}^{n} X_i$ is the best unbiased estimator i.e. UMVUE

2. Let $X_1, \dots X_n$ be i.i.d $N(\mu, \sigma^2)$.
Find the best unbiased estimator of σ^p, where p is a known positive constant, not necessarily an integer.

Sol^n: Let us layout a sketch.

If T is a complete sufficient statistic for θ, then $g(T)$ is complete sufficient for $g(\theta)$ for bijective g.

Since $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ is complete for σ^2,

by Lehman-Scheffé theorem, we need to find an unbiased estimator U for σ^p that is a function of S^2.

We know that $E[Y^r] = 2^r \frac{\Gamma(\frac{k}{2} + r)}{\Gamma(\frac{k}{2})}$ if $Y \sim \chi^2_k$

chi-squared with k degrees of freedom.

Also $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

We can probably look for a function involving $(S^2)^{P/2}$.

Let $Y = \frac{(n-1)S^2}{\sigma^2}$. Then $Y \sim \chi^2_{n-1}$

$$(S^2)^{P/2} = \left(\frac{\sigma^2 Y}{n-1}\right)^{P/2} = \frac{\sigma^P Y^{P/2}}{(n-1)^{P/2}}$$

$$\Rightarrow E[(S^2)^{P/2}] = \frac{\sigma^P}{(n-1)^{P/2}} E[Y^{P/2}]$$

$$E[Y^{P/2}] = 2^{P/2} \frac{\Gamma\left(\frac{n-1}{2} + \frac{P}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$\therefore E[(S^2)^{P/2}] = \frac{\sigma^P}{(n-1)^{P/2}} 2^{P/2} \frac{\Gamma\left(\frac{n-1}{2} + \frac{P}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$= \sigma^P \left(\frac{1}{n-1}\right)^{P/2} 2^{P/2} \frac{\Gamma\left(\frac{n-1+P}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

Now if we set $Z = \dfrac{(S^2)^{P/2}}{\left(\frac{1}{n-1}\right)^{P/2} 2^{P/2} \frac{\Gamma\left(\frac{n-1+P}{2}\right)}{\Gamma\left(\frac{n-1}{n}\right)}}$

This has expectation equal to σ^P.

i.e. $E[Z] = \sigma^P$ & this is UMVUE as it is a function of S^2 which is a complete statistic for σ^2.

Problem 2 (Loss function optimality, 20pts)

1. Show that if X is a continuous random variable, then $\min_a E|X-a| = \mathbb{E}|X-m|$ where m is the median of X.

Sol^n: The median 'm' is that value where the cumulative distribution function becomes $\frac{1}{2}$ i.e. $F(m) = \frac{1}{2}$.

Let $g(a) = E|X-a| = \int_{-\infty}^{\infty} |x-a| f(x)\, dx$

where $f(x)$ = probability density function & $F'(x) = f(x)$.

$$g(a) = \int_{-\infty}^{a} |x-a| f(x)\, dx + \int_{a}^{-\infty} |x-a| f(x)\, dx.$$

$$|x-a| = \begin{cases} a-x\, ; & x < a \\ x-a\, ; & x > a. \end{cases}$$

$$\therefore g(a) = \int_{-\infty}^{a} (a-x) f(x)\, dx + \int_{a}^{+\infty} (x-a) f(x)\, dx$$

At any minimum of $g(a)$, $g'(a)$ must change sign from $-$ve to $+$ve going through zero. i.e. $g'(a) = 0$

$$g'(a) = \frac{d}{da} \int_{-\infty}^{a} (a-x) f(x)\, dx + \int_{a}^{\infty} (x-a) f(x)\, dx$$

$$= \int_{-\infty}^{a} (+1) f(x)\, dx + \int_{a}^{+\infty} (-1) f(x)\, dx$$

$$= \int_{-\infty}^{a} f(x)\, dx - \int_{a}^{+\infty} f(x)\, dx$$

$\Rightarrow g'(a) = F(a) - (1 - F(a)) = 2F(a) - 1$

Setting $g'(a) = 0$, $2F(a) - 1 = 0 \Rightarrow F(a) = \frac{1}{2} \Rightarrow a = m$.

Q.E.D

Q2 Let $X_1 \dots X_n$ be a random sample from $N(\theta, \sigma^2)$ where σ^2 is known. Consider estimating θ using squared error loss. Let $\pi(\theta)$ be a $N(\mu, \tau^2)$ prior distribution on θ and δ^{π} be the Bayes estimator of θ. Prove the following holds:

a) For any constant a, b, the estimator $\delta(X) = a\bar{X} + b$ has risk function $R(\theta, \delta) = a^2 \frac{\sigma^2}{n} + (b - (1-a)\theta)^2$

Soln :-
$$R(\theta, \delta) = E_X[(\delta(X) - \theta)^2]$$
$$= E_X[(a\bar{X} + b - \theta)^2]$$
$$= E_X[\{a(\bar{X} - \theta) + b + (a-1)\theta\}^2]$$
$$= E_X[a^2(\bar{X} - \theta)^2 + \{b + (a-1)\theta\}^2 + 2a(\bar{X} - \theta)(b + (a-1)\theta)]$$
$$= a^2 E[(\bar{X} - \theta)^2] + [b + (a-1)\theta]^2 + 2a[b + (a-1)\theta]\,\underbrace{E[\bar{X} - \theta]}_{0}$$
$$= a^2 \operatorname{Var}(\bar{X}) + [b + (a-1)\theta]^2 + 0$$
$$= a^2 \frac{\sigma^2}{n} + [b + (a-1)\theta]^2$$

b) Let $\eta = \dfrac{\sigma^2}{n\tau^2 + \sigma^2}$. The risk function for the Bayes estimator is $R(\theta, \delta^\pi) = (1-\eta)^2 \dfrac{\sigma^2}{n} + \eta^2 (\theta - \mu)^2$

Sol^n $R(\theta, \delta^\pi) = \mathbb{E}\left[(\delta^\pi(x) - \theta)^2\right]$

$\pi(\theta \mid \bar{x}) \propto L(\theta; \bar{x})\, \pi(\theta)$

$$\propto \exp\left\{ \frac{-n}{2\sigma^2}(\bar{x} - \theta)^2 - \frac{1}{2\tau^2}(\theta - \mu)^2 \right\}$$

$$\propto \exp\left\{ \frac{-1}{2}\left(\frac{+n}{\sigma^2}(\bar{x} - \theta)^2 + \frac{1}{\tau^2}(\theta - \mu)^2 \right) \right\}$$

$$\propto \exp\left\{ \frac{-1}{2}\left(\frac{n}{\sigma^2}(\theta^2 - 2\bar{x}\theta) + \frac{1}{\tau^2}(\theta^2 - 2\mu\theta) \right) \right\}$$

$$\propto \exp\left\{ \frac{-1}{2}\left[\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2} \right)\theta^2 - 2\left(\frac{n}{\sigma^2}\bar{x} + \frac{\mu}{\tau^2} \right)\theta \right] \right\}$$

$$\theta_{post} = \frac{(n/\sigma^2)\bar{x} + \frac{\mu}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

$$\sigma_{post} = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}.$$

$$\therefore \delta^\pi(x) = \theta_{post} = \frac{\frac{n}{\sigma^2}\bar{x} + \frac{\mu}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

$$= \frac{n/\sigma^2}{n/\sigma^2 + \frac{1}{\tau^2}}\bar{x} + \frac{1/\tau^2}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}\mu = (1-\eta)\bar{x} + \eta\mu$$

$$\delta^{\pi}(x) = (1-\eta)\bar{x} + \eta\mu$$

$$R(\theta, \delta^{\pi}) = E[\{(1-\eta)\bar{x} + \eta\mu - \theta\}^2] = E[(\delta^{\pi} - \theta)^2]$$

$$= E[\{(1-\eta)(\bar{x}-\theta) + \eta(\mu-\theta)\}^2]$$

$$= E[(1-\eta)^2(\bar{x}-\theta)^2 + 2\eta(1-\eta)(\bar{x}-\theta)(\mu-\theta) + \eta^2(\mu-\theta)^2]$$

$$= (1-\eta)^2 E[(\bar{x}-\theta)^2] + \eta^2(\mu-\theta)^2 + 2\eta(1-\eta)(\mu-\theta)\underbrace{E[\bar{x}-\theta]}_{0}$$

$$= (1-\eta)^2\frac{\sigma^2}{n} + \eta^2(\mu-\theta)^2$$

<u>Q.E.D</u>

c) The Bayes risk of the Bayes estimator is

$$B(\pi, \delta^{\pi}) = \tau^2\eta.$$

<u>Sol</u>

Solⁿ $B(\pi, \delta^{\pi}) = \int R(\theta, \delta)\, d\pi(\theta)$

where $R(\theta, \delta) = E_{\theta}[(\delta(x) - \theta)^2]$

$$B(\pi, \delta^{\pi}) = \int (1-\eta)^2\frac{\sigma^2}{n} + \eta^2(\mu-\theta)^2$$

$$= (1-\eta)^2\frac{\sigma^2}{n} + \eta^2\underbrace{\int(\theta-\mu)^2\, d\pi(\theta)}_{\tau^2 \text{ as } \theta\sim N(\mu,\tau^2)}$$

$$= (1-\eta)^2\frac{\sigma^2}{n} + \tau^2\eta^2$$

$$B(\pi, \delta^{\pi}) = \underline{(1-\eta)^2 \frac{\sigma^2}{n} + \tau^2 \eta^2}$$

$\downarrow$

on simplification, this give us $\tau^2 \eta$.

$$\frac{\sigma^2}{n}(1-\eta)^2 = \frac{n^2 \tau^2}{(n\tau^2+\sigma^2)^2} \cdot \frac{\sigma^2}{n}$$

$$= \frac{n\tau^2 \cdot \sigma^2}{(n\tau^2+\sigma^2)^2}$$

Problem 3 : (EM algorithm, 20 pts)

Let $X_1 \ldots X_n$ be i.i.d exponential random variables with rate λ (i.e. pdf $f(x,\lambda) = \lambda e^{-\lambda x}$ for $x>0$)

For each $i = 1, \ldots, n$, define the indicator

$$Y_i = \mathbb{I}\{X_i > c_i\}$$

where $c_1, \ldots, c_n > 0$ are known constants.

a) Derive the EM recursion to compute the MLE of λ based on $Y_1 \ldots Y_n$.

b) Suppose $n=3$ and we observe $Y_1 = 1, Y_2 = 1, Y_3 = 0$ with thresholds $c_1 = 1, c_2 = 2, c_3 = 3$. If our initial guess is $\hat{\lambda}_0 = 1$, compute the first two EM iterates $\hat{\lambda}_1$ & $\hat{\lambda}_2$.

Soln a)

$$\underbrace{l_{complete}}_{\text{log-likelihood}}(\lambda; X_1 \ldots X_n) = \sum_{i=1}^{n} [\ln(\lambda) - \lambda X_i]$$

$$= n\ln(\lambda) - \lambda \sum_{i=1}^{n} X_i$$

E-step

$$Q(\lambda \mid \lambda^{(k)}) = E[\, l_{complete}(\lambda; X_1, \ldots X_n) \mid Y_1 \ldots Y_n, \lambda^{(k)}]$$

$$= E[\, n\ln(\lambda) - \lambda \sum X_i \mid Y_1 \ldots Y_n, \lambda^{(k)}]$$

$$= E[n\ln(\lambda) \mid Y_1 \ldots Y_n, \lambda^{(k)}]$$

$$- \lambda \underbrace{E[\sum X_i \mid Y_1 \ldots Y_n, \lambda^{(k)}]}_{S^{(k)}}$$

If $Y_i = \mathbb{1}(X_i > c_i)$

$$X_i \mid X_i > c_i \overset{d}{=} c_i + Z \quad \text{where } Z \sim \text{Exp}(\lambda^{(k)})$$

$$\therefore E[X_i \mid Y_i = 1, \lambda^{(k)}] = E[X_i \mid X_i > c_i, \lambda^{(k)}] = c_i + \frac{1}{\lambda^{(k)}}$$

If $Y_i = 0$ $(X_i \le c_i)$ then X_i is truncated to $[0, c_i]$

$$f_{X_i \mid X_i \le c_i}(x) = \frac{\lambda^{(k)} e^{-\lambda^{(k)} x}}{1 - e^{-\lambda^{(k)} c_i}}, \quad 0 \le x \le c_i$$

$$\therefore E[X_i \mid X_i \le c_i, \lambda^{(k)}] = \frac{1}{\lambda^{(k)}} - \frac{c_i e^{-\lambda^{(k)} c_i}}{1 - e^{-\lambda^{(k)} c_i}}$$

$$\therefore E[X_i \mid Y_i, \lambda^{(k)}] = \begin{cases} c_i + \frac{1}{\lambda^{(k)}} & ; \quad Y_i = 1 \\ \frac{1}{\lambda^{(k)}} - \frac{c_i e^{-\lambda^{(k)} c_i}}{1 - e^{-\lambda^{(k)} c_i}} & , \quad Y_i = 0 \end{cases}$$

Summing over $i = 1, \ldots, n$

$$\hat{S}^{(k)} = \sum E[X_i \mid Y_i \, \lambda^{(k)}]$$

M-Step $\quad Q(\lambda \mid \lambda^{(k)}) = n \ln(\lambda) - \lambda \hat{S}^{(k)}$

To maximize $\frac{d}{d\lambda} Q(\lambda \mid \lambda^{(k)}) = \frac{n}{\lambda} - \hat{S}^{(k)} = 0$

$$\Rightarrow \quad \lambda = \frac{n}{\hat{S}^{(k)}} = \frac{n}{\sum_{i=1}^{n} E[X_i \mid Y_i, \lambda^{(k)}]}$$

Part b) $n=3$, $Y_1=1$, $Y_2=1$, $Y_3=0$

$c_1=1$, $c_2=2$, $c_3=3$

$$E[x_i \mid Y_i=1,\lambda] = \begin{cases} c_i + \frac{1}{\lambda} & \text{if } Y_i=1 \\ \frac{1}{\lambda} - \frac{c_i e^{-\lambda c_i}}{1-e^{-\lambda c_i}}, & \text{otherwise.} \end{cases}$$

E-Step

$$\hat{S}^{(k)} = \sum_{i=1}^{3} E[x_i \mid Y_i, \lambda^{(k)}]$$

M-Step

$$\lambda^{(k+1)} = \frac{3}{\hat{S}^{(k)}}$$

Iteration 1

E-Step $\hat{S}^{(0)} = 2+3+0.842854$

$\simeq 5.842$

M-Step $\lambda^{(1)} = \frac{3}{\hat{S}^{(0)}} \approx 0.513.$

Iteration - 2

E-Step $\hat{S}^{(1)} = 2.949 + 3.949 + 1.132 \approx 8.03$

M-Step $\lambda^{(2)} = \frac{3}{\hat{S}^{(1)}} = \frac{3}{8.03} \approx 0.373$