# Tutorial - 3

Anurag Maurya

CSE, IITB

28 September 2020

# Q-1

- Consider a permutation of the first $n$ positive integers, generated uniformly randomly (i.e. each of the $n!$ different permutations are equally likely). The ordered pair $(i, j)$ in the permutation is called an inversion if $i < j$ but $j$ precedes $i$ (i.e. occurs earlier than $i$) in the permutation. Determine the expected number of inversions in a uniformly randomly generated permutation of the first n positive integers.

# Q-1

- Consider a permutation of the first $n$ positive integers, generated uniformly randomly (i.e. each of the $n!$ different permutations are equally likely). The ordered pair $(i, j)$ in the permutation is called an inversion if $i < j$ but $j$ precedes $i$ (i.e. occurs earlier than $i$) in the permutation. Determine the expected number of inversions in a uniformly randomly generated permutation of the first n positive integers.
  - Hints : Explicit PMF is difficult to compute(Closed Form)

# Q-1

- Consider a permutation of the first $n$ positive integers, generated uniformly randomly (i.e. each of the $n!$ different permutations are equally likely). The ordered pair $(i, j)$ in the permutation is called an inversion if $i < j$ but $j$ precedes $i$ (i.e. occurs earlier than $i$) in the permutation. Determine the expected number of inversions in a uniformly randomly generated permutation of the first n positive integers.
    - Hints : Explicit PMF is difficult to compute(Closed Form)
    - Hints : Is there a anything we can say about the symmetry of permutations.

# Q-1

- Consider a permutation of the first $n$ positive integers, generated uniformly randomly (i.e. each of the $n!$ different permutations are equally likely). The ordered pair $(i, j)$ in the permutation is called an inversion if $i < j$ but $j$ precedes $i$ (i.e. occurs earlier than $i$) in the permutation. Determine the expected number of inversions in a uniformly randomly generated permutation of the first n positive integers.
  - Hints : Explicit PMF is difficult to compute(Closed Form)
  - Hints : Is there a anything we can say about the symmetry of permutations.
  - Answer : Model inversion between numbers $i$ and $j$ as a r.v. $I_{ij}$

# Q-1 - Solution

- For pair $(i, j)$, let $I_{ij} = 1$ if $(i, j)$ is an inversion and 0 otherwise. And let $X$ be the total number of inversion pairs, Then

$$X = \sum_{1 \leq i < j \leq n} I_{ij}$$

- $\mathbb{E}(I_{ij}) = 1 \cdot P(I_{ij} = 1) + 0 \cdot P(I_{ij} = 0)$
  - $P(I_{ij} = 1) = P(I_{ij} = 0) = 1/2$
  - Since there are equal number of permutations where $i$ succeeds or precedes $j$ for any pair $(i, j)$.
- $\mathbb{E}(I_{ij}) = 1/2$
- Total possible pairs $= \binom{n}{2}$
- $\mathbb{E}(X) = \binom{n}{2} \cdot 1/2$
- Using linearity of $\mathbb{E}$ and independence of $I_{ij}$

# Q-2

- Consider independently drawn sample values $x_1, x_2, ..., x_n$, each from Poisson$(\lambda/n)$ where $n$ is known. What is the maximum likelihood estimate for $\lambda$? Derive the bias, variance, MSE of this estimator. Is this a consistent estimator? Why (not)?
  - (There is a physical significance to this question, even though one needn't understand it to answer the question. The noise in an image pixel is typically Poisson in nature. The values x1, x2, ..., xn correspond to n images of the same scene acquired in quick succession with acquisition time T /n per image, instead of acquiring one image in time T.)
  - Recall for a Poission$(\lambda)$ distribution,

$$Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

# Q-2 - Solution

- $x_i \sim Poission(\lambda/n)$
- Likelihood$(\lambda) = \prod_{1 \leq i \leq n} \dfrac{(\lambda/n)^{x_i} e^{-\lambda/n}}{x_i!}$, Therefore

$$NLL(\lambda) = \sum_{i=1}^{n} \frac{\lambda}{n} - x_i log(\frac{\lambda}{n}) + log(x_i!)$$

$$\frac{\partial NLL}{\partial \lambda} = \sum_{i=1}^{n} \frac{1}{n} - \sum_{i=1}^{n} \frac{x_i}{\lambda} = 0$$

$$\hat{\lambda} = \sum_{i=1}^{n} x_i$$

# Q-2 - Solution

$$\hat{\lambda} = \sum_{i=1}^{n} x_i$$

- $\mathbb{E}(\hat{\lambda}) = \frac{\lambda}{n} \cdot n = \lambda$
  - Thus, it's an unbiased estimator
- $\text{Var}(\hat{\lambda}) = \frac{\lambda}{n} \cdot n = \lambda$ (i.i.d random variables)
- $MSE = bias^2 + Var = \lambda$
  - The $MSE$ does not decrease with $n$, this is not a consistent estimator.

# Q-3

If $X \sim N(0,1)$, then prove that $P(|X| \geq u) \leq \sqrt{2/\pi}\dfrac{e^{-u^2/2}}{u}$ for all $u>0$. How does this bound compare with that given by Chebyshev's inequality?

# Q-3 - Solution

- $P(X \geq t) = P(X \leq -t)$, Thus $P(|X| \geq t) = 2P(X \geq t)$
  - By symmetry of gaussian around mean(0)

$$
\begin{aligned}
P(X \geq t) &= \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-x^2/2} dx \\
&\leq \frac{1}{\sqrt{2\pi}} \int_t^{\infty} \frac{x}{t} e^{-x^2/2} dx && as x \geq t \\
&= \frac{1}{t\sqrt{2\pi}} \int_{t^2/2}^{\infty} e^{-u} du && u = x^2/2, du = xdx \\
&= \frac{1}{t\sqrt{2\pi}} \left[ \frac{e^{-u}}{-1} \right]_{t^2/2}^{\infty} = \frac{e^{-t^2/2}}{t\sqrt{2\pi}}
\end{aligned}
$$

# Q-3 - Solution

- Thus $P(|X| \geq t) = \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}$
- Chebychev's inequality : $P(|Z| \geq t) \leq t^{-2}$
- Now $e^{-t^2/2}$ decreases faster than $t^{-2}$, this means the bound in this question is tighter than the one predicted by Chebychev's inequality.

# Q-4

Question : Consider $n$ values $\{x_i\}_{i=1}^{n}$ drawn independently from a Laplacian distribution with mean 0 and parameter $\sigma$. The probability density for a Laplacian random variable $X$ is given by $f_X(x) = \dfrac{1}{2\sigma} e^{-|x|/\sigma}$ (note the absolute value in the exponent). Given $\{x_i\}_{i=1}^{n}$, derive the maximum likelihood estimate for $\sigma$, as well as its bias, variance, MSE.

# Q-4 - Solution

- MLE for $\sigma$ is

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|X_i|$$

- Now $\mathbb{E}(|X|) = \int_{-\infty}^{\infty} \dfrac{|x|e^{-|x|/\sigma}}{2\sigma} dx$

$$\begin{aligned}
&= 2\int_0^\infty \frac{x}{2\sigma}e^{-x/\sigma}dx && \text{(using symmetry)}\\
&= \sigma\int_0^\infty \frac{x}{\sigma^2}e^{-x/\sigma}dx \\
&= \sigma\int_0^\infty ye^{-y}dy && (y = x/\sigma,\ dy = dx/\sigma)\\
&= \sigma\Big(\big[-ye^{-y}\big]_0^\infty - \int_0^\infty -e^{-y}dy\Big) = \sigma && u = y, dv = e^{-y}dy
\end{aligned}$$

Integration by Parts

# Q-4 - Solution

- Now $\mathbb{E}(|X|^2) = \int_{-\infty}^{\infty} \frac{x^2 e^{-|x|/\sigma}}{2\sigma} dx$

$$
\begin{aligned}
&= 2\int_0^\infty \frac{x^2}{2\sigma} e^{-x/\sigma} dx && \text{(using symmetry)} \\
&= \sigma^2 \int_0^\infty \frac{x}{\sigma^3} e^{-x/\sigma} dx && \\
&= \sigma^2 \int_0^\infty y^2 e^{-y} dy && (y = x/\sigma,\ dy = dx/\sigma) \\
&= 2\sigma^2 && \text{(Using Integration by parts)}
\end{aligned}
$$

# Q-4 - Solution

- Now $\mathbb{E}(\hat{\sigma}) = \frac{1}{n} n \cdot E(|X_i|) = \sigma$
    - i.i.d random variables
- $Var(\hat{\sigma}) = \frac{1}{n^2} n \cdot Var(X_i) = \sigma^2/n$
- $MSE = Var + bias^2 = \sigma^2/n$

# Q-5

Question : In this problem, we will derive higher order moments of specific random variables in a new way.

- Consider $X \sim N(\mu, \sigma^2)$ . Then prove that $E[g(X)(X-\mu)] = \sigma^2 E[g^{'}(X)]$ where $g$ is a differentiable function such that $E[|g^{'}(X)|] < \infty, |g(x)| < \infty$. Use this to derive an expression for $E[X^3]$ in terms of $\mu$ and $\sigma^2$. Do not use any other method (eg: MGFs) to derive $E[X^3]$.
- Consider $X \sim$Poisson$(\lambda)$. Then prove that $E[\lambda g(X)] = E[Xg(X-1)]$ where g is a function such that $-\infty < E[g(X)] < \infty, -\infty < g(-1) < \infty$. Use this to derive an expression for $E[X^3]$ assuming known expressions for $E[X]$, $E[X^2]$. Do not use any other method (eg: MGFs) to derive $E[X^3]$.

# Q-5 - Solution

(a)

$$E[g(X)(X-\mu)] = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} g(x)(x-\mu)e^{\dfrac{-(x-\mu)^2}{2\sigma^2}}dx$$

$$(u = g(x), dv = (x-\mu)e^{\dfrac{-(x-\mu)^2}{2\sigma^2}} \text{,Integration By Parts})$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left[-\sigma^2 g(x)e^{\dfrac{-(x-\mu)^2}{2\sigma^2}}\right]_{-\infty}^{\infty} + \sigma^2\int_{-\infty}^{\infty}\frac{g^{'}(x)}{\sqrt{2\pi}}e^{\dfrac{-(x-\mu)^2}{2\sigma^2}}dx$$

(First term is 0, Because $|g(x)| < \infty$)

$$= \sigma^2 E[g^{'}(X)]$$

# Q-5 - Solution

(a) - Contd.

$$
\begin{aligned}
E[X^3] &= E[X^2(X-\mu+\mu)] \\
&= E[X^2(X-\mu)] + \mu E[X^2] \\
&= \sigma^2 E(2X) + \mu(\mu^2+\sigma^2) \qquad \text{using } g(x)=x^2, g^{'}(x)=2x \\
&= 3\mu\sigma^2 + \mu^3
\end{aligned}
$$

# Q-5 - Solution

(b)

$$
\begin{aligned}
E[\lambda g(X)] &= \sum_{x=0}^{\infty} \lambda g(x) \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} \lambda g(x) \frac{e^{-\lambda}\lambda^x}{x!} \cdot \frac{x+1}{x+1} \\
&= \sum_{x=0}^{\infty} (x+1) g(x) \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!} \\
&= \sum_{y=1}^{\infty} y g(y-1) \frac{e^{-\lambda}\lambda^y}{y!} && \text{using } y = x+1 \\
&= E[Xg(X-1)]
\end{aligned}
$$

# Q-5 - Solution

(b) - Contd.

$$\begin{aligned} E[\lambda X^2] &= E[X(X-1)^2] && \text{using } g(x) = x^2 \\ &= E[X^3 - 2X^2 + X] \\ E[X^3] &= E[\lambda X^2] + 2E[X^2] - E[X] \\ &= \lambda^3 + 3\lambda^2 + \lambda \end{aligned}$$