Q1) MLE (MAXIMUM LIKELIHOOD ESTIMATOR) OF PARETO **(Total 5 points)**

Let $X_1, X_2, \ldots, X_n$ be i.i.d. as Pareto(θ), with the pdf of Pareto(θ) being $f(x) = \frac{\theta \cdot c^\theta}{x^{\theta+1}}$, where c > 0 is some given constant and assume that θ ≥ 1 and x ≥ c. Find the MLE of θ.

$$L(\theta) = \prod_{i=1}^{n} f(X_i | \theta) = \prod_{i=1}^{n} \frac{\theta \cdot c^\theta}{X_i^{\theta+1}}$$

$$\ell(\theta) = \log(L(\theta)) = \sum \log\left(\frac{\theta \cdot c^\theta}{X_i^{\theta+1}}\right)$$

log likelihood

$$= \sum (\log\theta + \theta \log c - (\theta+1)\log(X_i))$$

$$= n\log\theta + n\theta\log c - (\theta+1)\sum\log(X_i)$$

For MLE, $\frac{d\ell(\theta)}{d\theta} = 0$

$$\Rightarrow \frac{d}{d\theta}\left(n\log\theta + n\theta\log c - (\theta+1)\sum\log(X_i)\right) = 0$$

$$\Rightarrow \frac{n}{\theta} + n\log c - \sum\log(X_i) = 0$$

$$\Rightarrow \frac{n}{\theta} = \sum\log(X_i) - n\log c$$

$$\Rightarrow \hat{\theta}_{MLE} = \frac{n}{(\sum\log(X_i)) - n\log c}$$

Q2) MME (METHOD OF MOMENTS ESTIMATOR) WITH DATA **(Total 3 points)**

Let $X = \begin{cases} 1 & \textit{with prob } \theta \\ 3 & \textit{otherwise} \end{cases}$, with θ being an unknown parameter. Let D = {1, 1, 1, 3} be drawn i.i.d. from X. Derive $\hat{\theta}$, the MME of θ, using D as the sample data. Clearly show all your steps.

K = 1 unknown parameter

$$\hat{\alpha}_1(\theta) = \frac{\sum X_i}{n} = \bar{X}$$

$$\alpha_1(\theta) = E[X] = 1 \cdot \theta + 3 \cdot (1-\theta) = 3 - 2\theta$$

For MME, $\hat{\alpha}_1 = \alpha_1(\hat{\theta})$

$$\Rightarrow \bar{X} = 3 - 2\hat{\theta}$$

$$\Rightarrow \boxed{\hat{\theta} = \frac{3 - \bar{X}}{2}}$$

2

For $D = \{1, 1, 1, 3\}$ $\quad \bar{X} = \frac{1+1+1+3}{4} = 3/2$

$$\therefore \hat{\theta} = \frac{3 - \frac{3}{2}}{2} = \frac{3}{4}$$

-①

Q3) TOY EXAMPLE FOR PERMUTATION TEST **(Total 5 points)**

Let X = {1, 7} and Y = {3}. The null hypothesis is that X & Y are from the same distribution. The threshold for Accept/Reject is p=0.4.

(a) Use the permutation test to compute the p-value and decide the hypothesis. Please show all steps for each permutation clearly. (4 points)

(b) Explain, in words, why the Permutation test works. Consider the case where X and Y are test scores of two classes with very different intellect levels. (1 point)

(a) $T_{obs} = |\bar{X} - \bar{Y}| = \left|\frac{1+7}{2} - 3\right| = |4-3| = 1$ } 1/2

Permutations

X	Y	T_i ($\lvert\bar{X}-\bar{Y}\rvert$)	$I(T_i > T_{obs})$
{1,7}	{3}	1	0
{1,3}	{7}	~~3~~ 5	1
{3,1}	{7}	5	1
{3,7}	{1}	4	1
{7,1}	{3}	1	0
{7,3}	{1}	4	1

3

$$\text{p-value} = \frac{\sum I(T_i > T_{obs})}{n} = \frac{4}{6} = \frac{2}{3}$$

1/2

Since p-value > 0.4, we can't Reject, so accept null. (X = Y)

① (b) If X & Y are very different, $T_{obs} = |\bar{X} - \bar{Y}|$ is very large. So, $I(T_i > T_{obs})$ will likely be close to 0. Therefore, we will reject $\because$ p-value $\to 0$.
so $X \neq Y$

Q4) POISSON-GAMMA CONJUGATE **(Total 5 points)**

Let $X_1, X_2, \dots, X_n$ be i.i.d. as Poisson(λ). Recall that the p.m.f. of Poisson(λ) is p(x) = $\frac{e^{-\lambda}.\lambda^x}{x!}$. Let the prior of λ be Gamma(α, β), where the p.d.f. of Gamma(α, β) is such that f(x) is proportional to $x^{\alpha-1}e^{-x\beta}$. Show that the posterior is also a Gamma and find its parameters. Show all steps clearly. Feel free to ignore the constants and conclude that the posterior is a Gamma if its form resembles that of f(x) above.

Bayesian

$$f(\lambda \mid \{X_1, \dots, X_n\}) \propto \underbrace{L(\lambda)}_{\text{likelihood}} \cdot \underbrace{f(\lambda)}_{\text{prior}}$$

$$= \prod_{i=1}^{n} \frac{e^{-\lambda} \cdot \lambda^{X_i}}{(X_i!)} \times \lambda^{\alpha-1} \cdot e^{-\lambda\beta}$$

$$= \frac{e^{-n\lambda} \cdot \lambda^{\Sigma X_i}}{\pi (X_i!)} \cdot \lambda^{\alpha-1} \cdot e^{-\lambda\beta}$$

$$\propto e^{-\lambda(n+\beta)} \cdot \lambda^{\alpha-1+\Sigma X_i}$$

$$\sim \underline{\underline{\text{Gamma}(\alpha + \Sigma X_i, n+\beta)}}$$

Q5) WALD'S TEST **(Total 8 points)**

You observe 36 successes in 100 flips. The null hypothesis is that the coin is unbiased.

(a) Use the Wald's test with sample mean as estimator to Reject/Accept the null at the 95% confidence level (that is, $1-\alpha = 0.95$ and $z_{\alpha/2}=1.96$). Use a calculator, if needed. Show all steps, including why the Wald's test is applicable here and the derivation of the standard error, with explanations. (4 points)

(b) What is the p-value? You can leave your answer in terms of ϕ, the cdf of the standard normal. Show the area represented by the p-value using a Normal diagram. 1 (2 points)

(c) In the above test, what is the probability of a type I error? Show the area represented by this probability in the diagram from part (b). 1 (2 points)

(a) 1 $\left\{ \because \right.$ estimator is sample mean $\left(\frac{\sum X_i}{n}\right)$, by CLT, it is Asym Normal. So, Wald's is applicable.

$$W = \left| \frac{\hat{\theta} - \theta_0}{se(\hat{\theta})} \right| \qquad \hat{\theta} = \frac{\sum X_i}{n} = \frac{36}{100} \qquad \theta_0 = 0.5$$

Lov, iid 1/2

$$se(\hat{\theta}) = \sqrt{Var(\hat{\theta})} = \sqrt{Var\left(\frac{\sum X_i}{n}\right)} \overset{LoV}{=} \sqrt{\frac{1}{n^2} \cdot \sum Var(X_i)}$$

$$\overset{iid}{=} \sqrt{\frac{1}{n^2} \times n \, Var(X_1)} = \sqrt{\frac{Var(X_1)}{n}} = \sqrt{\frac{p(1-p)}{n}}, \because X \sim Bern.$$

$$\hat{se}(\hat{\theta}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} = \sqrt{\frac{\frac{36}{100} \times \frac{64}{100}}{100}}$$

$$= 0.048$$

$$\therefore W = \left| \frac{0.36 - 0.5}{0.048} \right| = \left| \frac{0.14}{0.048} \right| = \frac{140}{48} = 2.9167$$

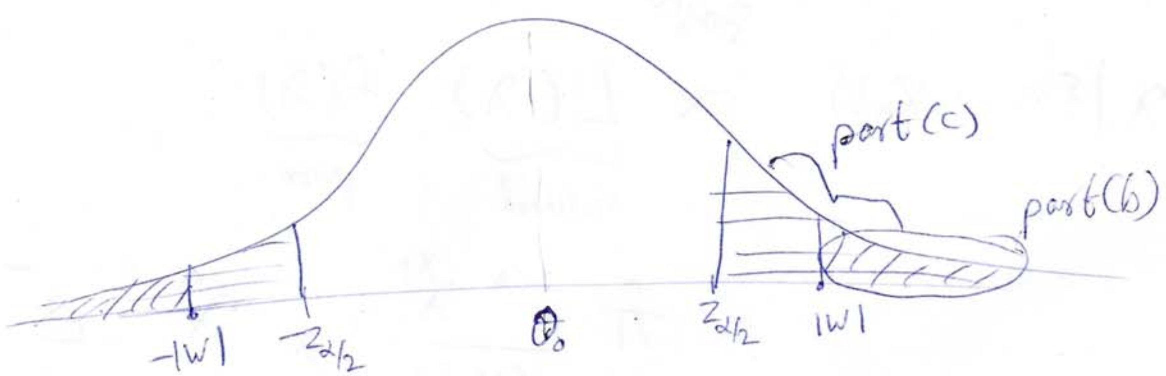
1/2 $\left\{ \because |W| > z_{\alpha/2} = 1.96 \right.$, we ~~can't~~ reject. ~~So, we accept.~~

(b) p-value = Pr(more extreme than |W|)

$$= \Pr(Z > |W|)$$

$$= 2(1 - \Pr(Z < |W|))$$

$$= 2(1 - \Phi(2.9167)).$$



(C) Type I error = incorrectly rejecting null

= Reject Null | Null is true

Pr(Type I) = Pr(Reject Null | Null is true)

$$= \Pr(|W| > Z_{\alpha/2})$$

$\because$ Null is true, $W \sim Z$

$$= 2 \cdot \Pr(Z > Z_{\alpha/2})$$

$$= 2(1 - \Phi(Z_{\alpha/2}))$$

$$= 2(1 - (1 - \alpha/2)) = 2(\alpha/2) = \underline{\underline{\alpha}}$$

Q6) SIMPLE LINEAR REGRESSION **(Total 4 points)**

Given data {(Y_1, X_1); (Y_2, X_2); …; (Y_n, X_n)}, consider a different form of simple linear regression where the estimate of Y_i is $\hat{Y}_i = \hat{\beta}.X_i$, with $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$. Note that there is no $\widehat{\beta_0}$ term here.

(a) Derive, from scratch, the OLS estimate of $\hat{\beta}$ that minimizes the SSE. (3 points)

(b) Find $\hat{\beta}$ for sample data {(3, 1); (7, 2); (9,3)}, where the pairs are (Y_i, X_i). (1 point)

(a) $SSE = \sum \hat{\varepsilon}_i^2 = \sum (Y_i - \hat{Y}_i)^2$

$= \sum (Y_i - (\hat{\beta} X_i))^2$

$$\frac{\partial SSE}{\partial \hat{\beta}} = 0 \Rightarrow \frac{\partial}{\partial \beta}\left(\sum (Y_i - \beta X_i)^2\right) = 0$$

$$\Rightarrow \sum \frac{\partial}{\partial \beta}(Y_i - \beta X_i)^2 = 0$$

$$\Rightarrow \sum 2.(Y_i - \beta X_i).(-X_i) = 0$$

$$\Rightarrow \left(\sum X_i Y_i\right) = \beta . \sum X_i^2 \Rightarrow \boxed{\hat{\beta}_{OLS} = \frac{\sum (X_i Y_i)}{\sum (X_i^2)}}$$

(b) $\sum (X_i Y_i) = 3 \cdot 1 + 7 \cdot 2 + 9 . 3 = 3 + 14 + 27 = 44$

$\sum X_i^2 = 1^2 + 2^2 + 3^2 = 1 + 4 + 9 = 14$

$\therefore \hat{\beta} = \frac{44}{14} = \frac{22}{7}$