Imperial College London

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)

May – June 2013

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science.

Probability & Statistics II

Date: Monday, 20 May 2013. Time: 2.00pm. Time allowed: 2 hours.

This paper has FOUR questions.

Candidates should start their solutions to each question in a new main answer book

Supplementary books may only be used after the relevant main book(s) are full.

Formula sheets are provided on pages 4 & 5.

- DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.
- Answer all the questions. Each question carries equal weight.
- Credit will be given for all questions attempted, but extra credit will be given for complete or nearly complete answers.
- Affix one of the labels provided to each answer book that you use, but DO NOT USE THE LABEL WITH YOUR NAME ON IT.
- Calculators may not be used.

1. Suppose $X_1, \ldots, X_n$ are iid random variables, each with PDF

$$f_X(x) = \frac{1}{2\sigma} \exp\left\{-\frac{|x-\mu|}{\sigma}\right\} \quad \text{for} \quad -\infty < x < \infty,$$

where $\sigma > 0$ and $-\infty < \mu < \infty$ are model parameters. For simplicity, assume that n is odd.

(a) Argue that

$$\frac{d}{d\mu}|x-\mu| = \begin{cases} -1 & \text{if } \mu < x \\ 1 & \text{if } \mu > x \end{cases}.$$

(The derivative does not exist for $\mu = x$, but you may ignore this.)
Derive the MLEs of μ and σ.
If you have trouble deriving the MLE of μ, just derive the MLE of σ for known μ.

(b) In this part suppose that μ is known to be zero.
Write down the simplified MLE of σ under this supposition. Call it $\widehat{\sigma}_0$.
Derive the distribution of $Y_i = |X_i|$. What named distribution is this?
State the distribution of $S = \sum_{i=1}^{n} |X_i|$. (No proof is needed if you use a standard result.)
Derive the sampling distribution of $n\widehat{\sigma}_0/\sigma$.

(c) Again assuming that $\mu = 0$, suppose we wish to construct a confidence interval for σ.
Propose a pivotal quantity for this purpose. State why your proposal is a pivot.
Use your pivot to construct a $100 \times (1-\alpha)\%$ confidence interval in terms of the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of a named distribution.
Can a shorter interval be constructed using the same pivot? (1-2 sentences are sufficient.)

2. (a) Suppose X and Y are independent standard normal RVs. Compute $\Pr(X^2 + Y^2 < 1)$.

(b) Let $X \sim \mathrm{N}(\mu, \sigma^2)$ and let the conditional distribution of Y given $X = x$ be $\mathrm{N}(\alpha+\beta x, \tau^2)$.
Compute $\mathrm{E}(Y)$, $\mathrm{Var}(Y)$, and the correlation of X and Y.

(c) Now let $X \sim \mathrm{N}(0, 1)$ and let the conditional distribution of Y given $X = x$ be $\mathrm{N}(x, 1)$.
Derive the conditional pdf of X given Y. State the name of this distribution.
Derive the marginal pdf of Y. Be sure to state the name of this distribution.

(d) Suppose $\mathcal{A}$ and $\mathcal{B}$ are both sigma algebras. Show that $\mathcal{A} \cap \mathcal{B}$ is also a sigma algebra.

3. Suppose $f_X(x)$ and $f_Y(y)$ are two probability density functions with the same support. Let

$$W = \frac{f_X(Y)}{f_Y(Y)} \qquad \text{and} \qquad T = \frac{Y f_X(Y)}{f_Y(Y)}.$$

(a) Compute $\mathrm{E}(T)$. Explain how this can be used to construct an unbiased estimator of $\mathrm{E}(X)$ from a random sample $(Y_1, \ldots, Y_n)$. Call your unbiased estimator S.

(b) Now suppose

$$f_X(x) = e^{-x} \text{ for } x > 0, \quad \text{and} \quad f_Y(y) = \gamma e^{-\gamma y} \text{ for } y > 0, \quad \text{with } \gamma > 0$$

Compute $\mathrm{Var}(T)$ and $\mathrm{Var}(S)$. Explain carefully how your answer depends on γ.
What value of γ minimizes the variances? Is this a sensible choice of γ?
What values of γ should absolutely be avoided when using S to estimate $\mathrm{E}(X)$? Why?

(c) Using the functional forms of f_X and f_Y given in part (b), derive the pdf of W. Be very careful about how your answer depends on γ.

4. Suppose $(X_1, \ldots, X_n)$ is a random sample from a Uniform $(0, \alpha)$ distribution, for some $\alpha > 0$. Let $T_n = 2\bar{X}_n = \frac{2}{n}\sum_{i=1}^n X_i$ and $M_n = \max(X_1, \ldots, X_n)$.

(a) State the definition of convergence in distribution, $X_n \xrightarrow{\mathcal{D}} X$.
State the definition of convergence in probability, $X_n \xrightarrow{\mathcal{P}} X$.

(b) Derive the sampling distribution of M_n.
Show that $M_n \xrightarrow{\mathcal{P}} \alpha$.

(c) Show that T_n is an unbiased estimator of α, but that M_n is a biased estimator of α.
Compute $\lim_{n\to\infty} \mathrm{bias}(M_n)$.

(d) This part deals with convergence in distribution.
What is the asymptotic (large n) distribution of $\sqrt{n}(T_n - \alpha)$? (Hint: use a theorem.)
Derive the asymptotic (large n) distribution of $n(\alpha - M_n)/\alpha$. (Prove your result.)

DISCRETE DISTRIBUTIONS

	RANGE $\mathbb{X}$	PARAMETERS	MASS FUNCTION f_X	CDF F_X	$\mathsf{E}_{f_X}[X]$	$\mathsf{Var}_{f_X}[X]$	MGF M_X
$Bernoulli(\theta)$	$\{0,1\}$	$\theta \in (0,1)$	$\theta^x(1-\theta)^{1-x}$		θ	$\theta(1-\theta)$	$1-\theta+\theta e^t$
$Binomial(n,\theta)$	$\{0,1,...,n\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$		$n\theta$	$n\theta(1-\theta)$	$(1-\theta+\theta e^t)^n$
$Poisson(\lambda)$	$\{0,1,2,...\}$	$\lambda \in \mathbb{R}^+$	$\dfrac{e^{-\lambda}\lambda^x}{x!}$		λ	λ	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,...\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\dfrac{1}{\theta}$	$\dfrac{(1-\theta)}{\theta^2}$	$\dfrac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(n,\theta)$	$\{n,n+1,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{x-1}{n-1}\theta^n(1-\theta)^{x-n}$		$\dfrac{n}{\theta}$	$\dfrac{n(1-\theta)}{\theta^2}$	$\left(\dfrac{\theta e^t}{1-e^t(1-\theta)}\right)^n$
or	$\{0,1,2,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n+x-1}{x}\theta^n(1-\theta)^x$		$\dfrac{n(1-\theta)}{\theta}$	$\dfrac{n(1-\theta)}{\theta^2}$	$\left(\dfrac{\theta}{1-e^t(1-\theta)}\right)^n$

For **CONTINUOUS** distributions (see over), define the **GAMMA FUNCTION**

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

and the LOCATION/SCALE transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = f_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} \qquad F_Y(y) = F_X\left(\frac{y-\mu}{\sigma}\right) \qquad M_Y(t) = e^{\mu t}M_X(\sigma t) \qquad \mathsf{E}_{f_Y}[Y] = \mu + \sigma\mathsf{E}_{f_X}[X] \qquad \mathsf{Var}_{f_Y}[Y] = \sigma^2\mathsf{Var}_{f_X}[X]$$

CONTINUOUS DISTRIBUTIONS

		PARAMS.	PDF	CDF	$\mathsf{E}_{f_X}[X]$	$\mathsf{Var}_{f_X}[X]$	MGF
	$\mathbb{X}$		f_X	F_X			M_X
$Uniform(\alpha,\beta)$ (stand. model $\alpha=0,\beta=1$)	(α,β)	$\alpha<\beta\in\mathbb{R}$	$\dfrac{1}{\beta-\alpha}$	$\dfrac{x-\alpha}{\beta-\alpha}$	$\dfrac{(\alpha+\beta)}{2}$	$\dfrac{(\beta-\alpha)^2}{12}$	$\dfrac{e^{\beta t}-e^{\alpha t}}{t(\beta-\alpha)}$
$Exponential(\lambda)$ (stand. model $\lambda=1$)	$\mathbb{R}^+$	$\lambda\in\mathbb{R}^+$	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	$\left(\dfrac{\lambda}{\lambda-t}\right)$
$Gamma(\alpha,\beta)$ (stand. model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$		$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$	$\left(\dfrac{\beta}{\beta-t}\right)^\alpha$
$Weibull(\alpha,\beta)$ (stand. model $\beta=1$)	$\mathbb{R}^+$	$\alpha,\beta\in\mathbb{R}^+$	$\alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}$	$1-e^{-\beta x^\alpha}$	$\dfrac{\Gamma(1+1/\alpha)}{\beta^{1/\alpha}}$	$\dfrac{\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma\left(1+\frac{1}{\alpha}\right)^2}{\beta^{2/\alpha}}$	
$Normal(\mu,\sigma^2)$ (stand. model $\mu=0,\sigma=1$)	$\mathbb{R}$	$\mu\in\mathbb{R},\sigma\in\mathbb{R}^+$	$\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$		μ	σ^2	$e^{\{\mu t+\sigma^2t^2/2\}}$
$Student(\nu)$	$\mathbb{R}$	$\nu\in\mathbb{R}^+$	$\dfrac{(\pi\nu)^{-\frac{1}{2}}\Gamma\left(\dfrac{\nu+1}{2}\right)}{\Gamma\left(\dfrac{\nu}{2}\right)\left\{1+\dfrac{x^2}{\nu}\right\}^{(\nu+1)/2}}$		0 (if $\nu>1$)	$\dfrac{\nu}{\nu-2}$ (if $\nu>2$)	
$Pareto(\theta,\alpha)$	$\mathbb{R}^+$	$\theta,\alpha\in\mathbb{R}^+$	$\dfrac{\alpha\theta^\alpha}{(\theta+x)^{\alpha+1}}$	$1-\left(\dfrac{\theta}{\theta+x}\right)^\alpha$	$\dfrac{\theta}{\alpha-1}$ (if $\alpha>1$)	$\dfrac{\alpha\theta^2}{(\alpha-1)(\alpha-2)}$ (if $\alpha>2$)	
$Beta(\alpha,\beta)$	$(0,1)$	$\alpha,\beta\in\mathbb{R}^+$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$		$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	

M2S1 FINAL EXAM SOLUTION 2012 - 2013.

1a) If $\mu < x$, $|x-\mu| = x-\mu$ and $\frac{d}{d\mu}|x-\mu| = -1$.

$\mu > x$, $|x-\mu| = \mu - x$ and $\frac{d}{d\mu}|x-\mu| = 1$.

1 mark

$$\ell(\mu,\sigma) = -n\log\sigma - \frac{1}{\sigma}\sum_{i=1}^{n}|x_i-\mu| + c$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{\sigma}\left[\sum_i I\{\mu > x_i\} - \sum_i I\{\mu < x_i\}\right] = 0$$

This quantity will be zero if an equal number of x_i are less than and greater than μ, i.e., if $\mu = \text{median}(x_1,\dots,x_n)$.

So $\hat{\mu}_{MLE} = \text{median}(x_1,\dots,x_n)$.

Although $\partial\ell/\partial\mu$ does not exist ~~for~~ at each $\mu = x_i$, ℓ is a continuous function of μ. ~~and~~ Further, ℓ is increasing for $\mu < \hat{\mu}_{MLE}$ and decreasing for $\mu > \hat{\mu}_{MLE}$, thus $\hat{\mu}_{MLE}$ is the maximizer of ℓ.

3 marks

$$\frac{\partial \ell}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|x_i-\mu| = 0 \Rightarrow \sigma_{MLE} = \frac{1}{n}\sum_{i=1}^{n}|x_i-\hat{\mu}_{MLE}|.$$

$$\frac{\partial^2 \ell}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_i|x_i-\mu|, \text{ evaluated at the MLE this is}$$

$$= \frac{n}{\hat{\sigma}^2_{MLE}} - 2\frac{n\hat{\sigma}_{MLE}}{\hat{\sigma}^3_{MLE}} = -\frac{n}{\hat{\sigma}_{MLE}} < 0 \Rightarrow \text{maximum}.$$

b) $\boxed{\hat{\sigma}_0 = \frac{1}{n}\sum_i|x_i|.}$ 1 mark

$$\underset{0<y<+\infty}{Y_i = |X_i|} \Longrightarrow X_i = \begin{cases} Y_i & \text{if } Y_i > 0 \\ -Y_i & \text{if } Y_i < 0 \end{cases} \qquad \left|\frac{dx}{dy}\right| = 1.$$

$$f_{Y_i}(y_i) = f_X(y_i) + f_X(-y_i) = \frac{1}{2\sigma}\exp\left\{-\frac{|y_i|}{\sigma}\right\} + \frac{1}{2\sigma}\exp\left\{-\frac{|-y_i|}{\sigma}\right\}$$

$$= \frac{1}{\sigma}\exp\left\{-\frac{y_i}{\sigma}\right\} \quad \text{for } 0 < y_i.$$

2 marks

$\boxed{\text{I.e., } Y_i \sim \exp\left(\frac{1}{\sigma}\right).}$ 1 mark

2 marks

We know sum of n iid exponentials is gamma (n, λ), where λ is the scale parameter of the exponentials. Thus $S \sim$ gamma $(n, \frac{1}{\sigma})$.

2 marks

$\frac{n\hat{\sigma}_0}{\sigma} = \frac{S}{\sigma} \sim$ Gamma $(n, 1)$, noting that the second parameter of a gamma dist'n is a scale parameter.

1 mark

c) $\boxed{\frac{n\hat{\sigma}_0}{\sigma} \text{ is an appropriate pivot.}}$ This quantity depends only on the random sample and σ and has a known, completely specified dist'n.

Let g_L be the $\frac{\alpha}{2}$ quantile of gamma $(n, 1)$
g_U be the $1-\frac{\alpha}{2}$ quantile of gamma $(n, 1)$.

3 marks

$$\Pr\left(g_L \le \frac{n\hat{\sigma}_0}{\sigma} \le g_U\right) = 100(1-\alpha)\%$$

Solving for σ,

$$\Pr\left(\frac{n\hat{\sigma}_0}{g_U} \le \sigma \le \frac{n\hat{\sigma}_0}{g_L}\right) = 100(1-\alpha)\%$$

So $\left(\frac{n\hat{\sigma}_0}{g_U}, \frac{n\hat{\sigma}_0}{g_L}\right)$ is a $100(1-\alpha)\%$ CI.

2 marks

Following the above calculation, any interval of the form

$$\left(\frac{n\hat{\sigma}_0}{g_U^\star}, \frac{n\hat{\sigma}_0}{g_L^\star}\right) = n\hat{\sigma}_0\left(\frac{1}{g_U^\star}, \frac{1}{g_L^\star}\right)$$

is a $100 \times (1-\alpha)\%$ confidence interval, just so long as $\Pr(g_L^\star < V < g_U^\star) = 100 \times (1-\alpha)\%$, where $V \sim$ GAMMA$(n, 1)$. Finding $g_L^\star$ and $g_U^\star$ satisfying this constraint, such that

$$\frac{1}{g_L^\star} - \frac{1}{g_U^\star} < \frac{1}{g_L} - \frac{1}{g_U}$$

results in a shorter interval.

An alternate solution is to argue that the minimum-width confidence interval at a given confidence level is a set of the form $\{x : f_X(x|\sigma^2) \ge c\}$ where c is a constant and f_X is the pdf of the pivot, in this case, $n\hat{\sigma}_0/\sigma$.

2a) $$\Pr(X^2+Y^2<1) = \iint_{x^2+y^2<1} \frac{1}{2\pi}\exp\left\{-\tfrac{1}{2}(x^2+y^2)\right\} = \int_0^{2\pi}\int_0^1 \frac{r}{2\pi} e^{-r^2/2}\,dr\,d\theta$$

$$= \int_0^{2\pi} \frac{1}{2\pi}\left[-e^{-r^2/2}\right]_0^1 d\theta = \left(1-e^{-1/2}\right)\int_0^{2\pi}\frac{1}{2\pi}\,d\theta = 1-e^{-1/2}.$$

Alternatively, $U = X^2+Y^2 \sim \chi^2_2$, so $U \sim \exp(\tfrac{1}{2})$ and $\Pr(U<1) = \int_0^1 \tfrac{1}{2}e^{-u/2}\,du = 1-e^{-1/2}$.

b) $E(Y) = \cancel{E[E(Y|X)]}\ E[E(Y|X)] = E[\alpha+\beta X] = \alpha+\beta\mu.$ ← 2 marks

$$\mathrm{Var}(Y) = E[\mathrm{Var}(Y|X)] + \mathrm{Var}[E(Y|X)]$$
$$= E[\tau^2] + \mathrm{Var}(\alpha+\beta X) = \tau^2+\beta^2\sigma^2$$

2 marks

$$E(XY) = E[E(XY|X)] = E[XE(Y|X)] = E(\alpha X+\beta X^2) = \alpha\mu+\beta(\mu^2+\sigma^2)$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \alpha\mu+\beta(\mu^2+\sigma^2) - (\alpha\mu+\beta\mu^2) = \beta\sigma^2$$

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \beta\sigma/\sqrt{\tau^2+\beta^2\sigma^2}.$$

c) $$f_{X|Y=y}(x|y) \propto f_{XY}(x,y) = \frac{1}{2\pi}\exp\left\{-\tfrac{1}{2}\left[x^2+(y-x)^2\right]\right\}$$

$$= \frac{1}{2\pi}\exp\left\{-\tfrac{1}{2}\left[x^2+y^2-2xy+x^2\right]\right\}$$

$$\propto \frac{1}{2\pi}\exp\left\{-\tfrac{1}{2}\left[2x^2-2xy\right]\right\} \propto \frac{1}{2\pi}\exp\left\{-\left(x-\tfrac{y}{2}\right)^2\right\}$$

$$\Rightarrow \boxed{X|Y \sim N\left(\tfrac{y}{2},\tfrac{1}{2}\right)}$$ 1 mark

2 marks $$f_Y(y) = \frac{f_{XY}(x,y)}{f_{X|Y=y}(x|y)} = \frac{\frac{1}{2\pi}\exp\left\{-\frac{1}{2}\left[x^2+y^2-2xy+x^2\right]\right\}}{\sqrt{\frac{2}{2\pi}}\exp\left\{-\left(x^2-xy+\frac{y^2}{4}\right)\right\}} = \frac{1}{2\sqrt{\pi}}\exp\left\{-\tfrac{1}{2}\left[\tfrac{y^2}{2}\right]\right\}$$

$$\Rightarrow \boxed{Y \sim N(0,2)}$$ 1 mark

d) Since $\mathcal{A}$ is a sigma algebra, we know

i) $\emptyset \in \mathcal{A}$

ii) if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$

iii) if $A_1, A_2, \ldots \in \mathcal{A}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

Likewise for $\mathcal{B}$.

Now

i) Because $\emptyset \in \mathcal{A}$ and $\emptyset \in \mathcal{B}$, $\emptyset \in \mathcal{A} \cap \mathcal{B}$

ii) If $A \in \mathcal{A} \cap \mathcal{B}$ then $A \in \mathcal{A}$ so $A^c \in \mathcal{A}$ and $A \in \mathcal{B}$ so $A^c \in \mathcal{B}$. Thus $A^c \in \mathcal{A} \cap \mathcal{B}$.

iii) If $A_1, A_2, \ldots \in \mathcal{A} \cap \mathcal{B}$ then $A_1, A_2, \ldots \in \mathcal{A}$ so $\bigcup_i A_i \in \mathcal{A}$ and $A_1, A_2, \ldots \in \mathcal{B}$ so $\bigcup_i A_i \in \mathcal{B}$. Thus $\bigcup_i A_i \in \mathcal{A} \cap \mathcal{B}$.

3 marks

3a) $E(T) = \int y \frac{f_X(y)}{f_Y(y)} f_Y(y)\, dy = \int y f_X(y)\, dy = E(X)$. 2 marks

Let

$$S = \frac{1}{n}\sum_{i=1}^{n} Y_i \frac{f_X(Y_i)}{f_Y(Y_i)}. \quad E(S) = \frac{1}{n}\sum_{i=1}^{n} E\left[Y_i \frac{f_X(Y_i)}{f_Y(Y_i)}\right] = E(X),$$

by the above calculation. Given a random sample of Y_i, we can estimate $E(X)$ using the estimator S. S is an unbiased estimator of $E(X)$ because $E(S) = E(X)$. (1) mark

b)

$$E(T^2) = \int y^2 \frac{f_X^2(y)}{f_Y^2(y)} f_Y(y)\, dy = \int_0^\infty \frac{1}{\gamma} y^2 e^{-y(2-\gamma)}\, dy$$

$$= \frac{\Gamma(3)}{\gamma(2-\gamma)^3} \quad \text{if } \gamma < 2.$$

$E(T^2)$ does not exist for $\gamma \geq 2$. 3 marks

$E(T) = E(X) = 1$ for the given $f_X(x)$.

$$Var(T) = \frac{\cancel{\Gamma(3)}\, 2}{\gamma(2-\gamma)^3} - 1 \quad \text{for } \gamma < 2$$

$$= +\infty \quad \text{for } \gamma \geq 2.$$

2 marks (and for noting Var does not exist for $\gamma \geq 2$)

$$Var(S) = \frac{Var(T)}{n} = \frac{2}{n\gamma(2-\gamma)^3} - \frac{1}{n} \quad \text{for } \gamma < 2$$

$$= +\infty \quad \text{for } \gamma \geq 2.$$

We can minimize the variances, by maximizing $\gamma(2-\gamma)^3$ or $\log\gamma + 3\log(2-\gamma)$. Differentiating, we find $\boxed{\gamma = 1}$ minimizes the variances. This is a sensible choice as it results is $T \sim X$. ← (1) mark

Values of $\gamma \geq 2$ result is estimators of $E(X)$ with infinite variance. For these values there is no guarantee $S \xrightarrow{?} E(X)$ since the law of large numbers does not apply. (1) mark

c) $W = \frac{1}{\gamma} e^{-Y(1-\gamma)}$ $\quad Y = \frac{-\log(\gamma W)}{1-\gamma}$ $\quad \left|\frac{dy}{dw}\right| = \frac{1}{w|1-\gamma|}$ for $\gamma \neq 1$.

(If $\gamma = 1$, $W \equiv 1$. So we may ignore this trivial case.)

Support of Y is $(0, \infty)$.

Support of W is $\begin{cases} (0, 1/\gamma) & \text{if } \gamma < 1 \\ 1 & \text{if } \gamma = 1 \\ (1/\gamma, \infty) & \text{if } \gamma > 1 \end{cases}$

$$f_W(w) = f_Y\left(\frac{-\log(\gamma\omega)}{1-\gamma}\right)\left|\frac{dy}{dw}\right| = \frac{\gamma}{|1-\gamma|\omega} \exp\left\{\frac{\gamma}{1-\gamma}\log(\gamma\omega)\right\}$$

$$= \frac{\gamma}{|1-\gamma|\omega}(\gamma\omega)^{\gamma/1-\gamma} = \frac{\sqrt[1-\gamma]{\gamma}}{1-\gamma}\,\omega^{\frac{2\gamma-1}{1-\gamma}} \text{ for } \begin{cases} 0 < \omega < 1/\gamma & \text{if } \gamma < 1 \\ \frac{1}{\gamma} < \omega < +\infty & \text{if } \gamma > 1. \end{cases}$$

TO CHECK YOUR ANSWER:

We know $E(W) = \int \frac{f_X(y)}{f_Y(y)} f_Y(y)\,dy = 1.$

Computing from the pdf.

$$E(W) = \frac{\gamma}{1-\gamma}\int_0^{1/\gamma}(\gamma\omega)^{\gamma/1-\gamma}\,d\omega = \frac{\gamma^{1/1-\gamma}}{1-\gamma}\left[(1-\gamma)\,\omega^{1/1-\gamma}\right]_0^{1/\gamma} = 1 \quad \text{if } \gamma < 1$$

$$= \frac{\gamma^{1/1-\gamma}}{\gamma-1}\int_{1/\gamma}^{\infty}\omega^{\gamma/1-\gamma}\,d\omega = \frac{\gamma^{1/1-\gamma}}{\gamma-1}\left[(1-\gamma)\,\omega^{1/1-\gamma}\right]_{1/\gamma}^{\infty} = 1 \quad \text{if } \gamma > 1$$

4a) We say $X_n \xrightarrow{\mathcal{D}} X$ if $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ at all points of continuity of F_X.

We say $X_n \xrightarrow{P} X$ if $\forall \epsilon > 0$ $\lim_{n\to\infty} \Pr(|X_n - X| < \epsilon) = 1$.

b) We note $M_n \le m$ if and only if $X_1, \dots, X_n \le m$.

$F_{M_n}(m) = \Pr(M_n \le m) = \prod_{i=1}^{n} \Pr(X_i \le m) = \left(\frac{m}{\alpha}\right)^n$ for $0 \le m \le \alpha$

$$\lim_{n\to\infty} F_{M_n}(m) = \lim_{n\to\infty} \left(\frac{m}{\alpha}\right)^n = 0 \quad \text{for } m < \alpha$$

$$= 1 \quad \text{for } m \ge \alpha$$

So $M_n \xrightarrow{\mathcal{D}} \alpha$. But convergence in dist'n to a constant is equivalent to convergence in probability to a constant. I.e., $M_n \xrightarrow{P} \alpha$.

c) $E(T_n) = \frac{2}{n} \sum_{i=1}^{n} E(X_i) = 2E(X_1) = 2\frac{\alpha}{2} = \alpha$.

$$f_{M_n}(m) = \frac{d}{dm} F_{M_n}(m) = n\left(\frac{m}{\alpha}\right)^{n-1} \frac{1}{\alpha} \quad \text{for } 0 \le m \le \alpha$$

$$E(M_n) = \int_0^\alpha \frac{n}{\alpha^n} m^n \, dm = \frac{n}{(n+1)\alpha^n} m^{n+1} \Big|_0^\alpha = \frac{n\alpha}{n+1} \ne \alpha$$

$$\text{Bias}(M_n) = E(M_n) - \alpha = \alpha\left(\frac{n}{n+1} - 1\right) = \frac{-\alpha}{n+1} \to 0 \quad \text{as } n \to \infty.$$

d) $E(X_1) = \frac{\alpha}{2}$, $\text{Var}(X_1) = \frac{\alpha^2}{12}$

So By the CLT $$\frac{\sqrt{n}\left(\bar{X}_n - \frac{\alpha}{2}\right)}{\alpha/\sqrt{12}} \xrightarrow{\mathcal{D}} N(0,1)$$

$$\frac{\sqrt{12}}{2\alpha} \sqrt{n}(T_n - \alpha) \xrightarrow{\mathcal{D}} N(0,1)$$

$$\sqrt{n}(T_n - \alpha) \xrightarrow{\mathcal{D}} N\left(0, \frac{\alpha^2}{3}\right).$$

3 marks

d) Let $U_n = n(\alpha - M_n)/\alpha$

$$F_{U_n}(u) = Pr(U_n \le u) = Pr\left(n(\alpha - M_n)/\alpha \le u\right) = Pr\left(M_n \geq \alpha\left(1-\frac{u}{n}\right)\right)$$
$$= 1 - F_{M_n}\left(\alpha\left(1-\frac{u}{n}\right)\right) = 1 - \left(1-\frac{u}{n}\right)^n$$

3 marks

$$\lim_{n\to\infty} F_{U_n}(u) = 1 - \lim_{n\to\infty}\left(1-\frac{u}{n}\right)^n = 1 - e^{-u}$$

I.e., $U_n \to$ Exponential (1).

Marks

FINAL

1. Suppose $X_1, \ldots, X_n$ are iid random variables, each with PDF

$$f_X(x) = \frac{1}{2\sigma} \exp\left\{-\frac{|\, x - \mu \,|}{\sigma}\right\} \quad \text{for} \quad -\infty < x < \infty,$$

where $\sigma > 0$ and $-\infty < \mu < \infty$ are model parameters. For simplicity, assume that n is odd.

(a) 1 Argue that

$$\frac{d}{d\mu}\,|\, x - \mu| = \begin{cases} -1 & \text{if } \mu < x \\ 1 & \text{if } \mu > x \end{cases}.$$

(The derivative does not exist for $\mu = x$, but you may ignore this.)

3 Derive the MLEs of μ and σ.

If you have trouble deriving the MLE of μ, just derive the MLE of σ for known μ.

(b) In this part suppose that μ is known to be zero.

1 Write down the simplified MLE of σ under this supposition. Call it $\widehat{\sigma}_0$.

2 Derive the distribution of $Y_i = |X_i|$. 1 What named distribution is this?

2 State the distribution of $S = \sum_{i=1}^{n} |X_i|$. (No proof is needed if you use a standard result.)

2 Derive the sampling distribution of $n\widehat{\sigma}_0/\sigma$.

(c) Again assuming that $\mu = 0$, suppose we wish to construct a confidence interval for σ.

1 Propose a pivotal quantity for this purpose. 2 State why your proposal is a pivot.

3 Use your pivot to construct a $100 \times (1 - \alpha)\%$ confidence interval in terms of the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of a named distribution.

2 Can a shorter interval be constructed using the same pivot? (1-2 sentences are sufficient.)

2. (a) Suppose X and Y are independent standard normal RVs. 3 Compute $\Pr(X^2 + Y^2 < 1)$.

(b) Let $X \sim \mathrm{N}(\mu, \sigma^2)$ and let the conditional distribution of Y given $X = x$ be $\mathrm{N}(\alpha + \beta x, \tau^2)$. Compute 2 $\mathrm{E}(Y)$, 2 $\mathrm{Var}(Y)$, and 3 the correlation of X and Y.

(c) Now let $X \sim \mathrm{N}(0, 1)$ and let the conditional distribution of Y given $X = x$ be $\mathrm{N}(x, 1)$.

3 Derive the conditional pdf of X given Y. 1 State the name of this distribution.

2 Derive the marginal pdf of Y. 1 Be sure to state the name of this distribution.

(d) Suppose $\mathcal{A}$ and $\mathcal{B}$ are both sigma algebras. 3 Show that $\mathcal{A} \cap \mathcal{B}$ is also a sigma algebra.