PhD Prelim Exam THEORY

(Majors)

Summer 2014
(Given on 7/17/14)

Part I

Suppose that given $\theta \in R^1$, $X \sim N(\theta, 1)$. Also assume $\theta \sim N(0, 1)$.

1. Show that the joint distribution of X and θ is bivariate normal. Identify its mean and covariance matrix.

2. What is the distribution of $(X - \theta)^2 + \theta^2$? Prove your assertion.

3. Prove that $\frac{X}{\theta}$ has the same distribution as $cW + d$ where W has a standard Cauchy distribution, and c and d are two constants. (Hint: you may use without proof the fact that if U and V are iid $N(0, 1)$, then $\frac{U}{V}$ has a standard Cauchy distribution.)

Part II

Consider a function $\delta(u) : R^1 \to R^1$ with the following three properties:

- $\delta(u) = 0$ if $u \leq 0$ or $u \geq 1$;
- $|\delta(u)| \leq 1$;
- $\int_0^1 \delta(u)du = 0$.

For $\mathbf{u} = (u_1, ..., u_n)^T \in R^n$, define a function $f(\mathbf{u})$: $R^n \to R^1$ by

$$f(\mathbf{u}) = \begin{cases} 1 - \prod_{i=1}^n \delta(u_i) & \text{if } 0 < u_i < 1 \text{ for all } i = 1, ..., n \\ 0 & \text{otherwise.} \end{cases}$$

4. Show that $f(\mathbf{u})$ is a density function.

5. Suppose a random variable $\mathbf{U} = (U_1, ..., U_n)^T$ has density $f(\mathbf{u})$.

   (a) Show that any subset with $n - 1$ elements from $\mathbf{U}$ are iid Uniform $(0, 1)$.

   (b) Give an example of a function $\delta(u)$ so that $U_1, ..., U_n$ are NOT independent.

6. For $\mathbf{u} \in R^n$, define $G(\mathbf{u}) = \int_{-\infty}^{u_1} \cdots \int_{-\infty}^{u_n} f(\mathbf{v})d\mathbf{v}$ with the density function $f(\mathbf{u})$ defined above. Suppose $F_i(x)$ is a distribution function on R^1 for $i = 1, ..., n$. Let $H(x_1, ..., x_n) = G(F_1(x_1), ..., F_n(x_n))$ be a function of $\mathbf{x} = (x_1, ..., x_n)^T \in R^n$.

   (a) Show that $H(x_1, ..., x_n)$ is a distribution function with $F_1(x_1)$, ..., $F_n(x_n)$ as its marginal distribution functions.

   (b) If a random variable $\mathbf{X} = (X_1, ..., X_n)^T$ has distribution function $H(x_1, ..., x_n)$, show that X_1, ..., X_{n-1} are independent.

Part III

Let $\mathbf{X} = (X_1, X_2, X_3)^T$. Suppose $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$ and the 3×3 covariance matrix $\boldsymbol{\Sigma}$ is positive definite. Let ρ_{ij} be the correlation between X_i and X_j for $i, j = 1, 2, 3$.

7. Prove that conditional on X_3 the random variables X_1 and X_2 are independent if only if $\rho_{12} = \rho_{13}\rho_{23}$.

Part IV

Suppose X_1, ..., X_n are iid $N(\theta, 1)$ for $\theta \in R^1$.

8. Find a UMVUE for $\Phi(x_0 - \theta)$ where $\Phi(\cdot)$ is the cdf of the $N(0, 1)$ distribution and x_0 is a fixed number. Argue carefully that your estimator is indeed UMVUE.

Part V

Let X_1,...,X_n be iid $N(\theta, \theta)$ with $\theta > 0$ (i.e. $E(X_i) = Var(X_i) = \theta$).

9. Find solutions of θ to the log likelihood equations.

10. Show that one of the solutions obtained in (9) (denoted by $\hat{\theta}_n$) is consistent, that is show $\hat{\theta}_n \xrightarrow{p} \theta$.

11. For the $\hat{\theta}_n$ obtained in (10), prove that $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \nu(\theta))$ and find $\nu(\theta)$ explicitly.

Basic Theory I Key   Statistics PhD – July 2014

Part I:

1. $f_{X|\theta}(x,\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}$ and $f_\theta(\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}$

So the joint p.d.f of (X,θ) is

$$f(x,\theta) = f_{X|\theta}(x,\theta) \times f_\theta(\theta) = \frac{1}{2\pi} e^{-\frac{(x-\theta)^2}{2} - \frac{\theta^2}{2}}$$
$$= \frac{1}{2\pi} e^{-\frac{1}{2}(x^2 - 2\theta x + 2\theta^2)}$$
$$= \frac{1}{2\pi} e^{-\frac{1}{2}\binom{x}{\theta}^T \Sigma^{-1} \binom{x}{\theta}}$$

where $\Sigma^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$. Then $\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

$\Rightarrow \binom{X}{\theta} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\right)$, a bivariate Normal.

2. $(x-\theta)^2 + \theta^2 = x^2 - 2\theta x + 2\theta^2 = \begin{bmatrix} x \\ \theta \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x \\ \theta \end{bmatrix}$

$$= \left\{\Sigma^{-\frac{1}{2}} \begin{bmatrix} x \\ \theta \end{bmatrix}\right\}^T \left\{\Sigma^{-\frac{1}{2}} \begin{bmatrix} x \\ \theta \end{bmatrix}\right\}$$

$\Sigma^{-\frac{1}{2}} \begin{bmatrix} x \\ \theta \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma^{-\frac{1}{2}} \Sigma \Sigma^{-\frac{1}{2}}\right) = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, I_2\right)$

So $(x-\theta)^2 + \theta^2 \sim \chi^2_{df=2}$, a χ^2-diseribution with degrees of freedom $= 2$.

3. Let $Y = aX + b\theta$. Because $\begin{bmatrix} X \\ \theta \end{bmatrix}$ is a bivariate Normal, so Y must be a Normal.

$\mathrm{cov}(Y, \theta) = \mathrm{cov}(aX + b\theta, \theta) = a\,\mathrm{cov}(X, \theta) + b\,V(\theta)$

$= a + b$

If we choose a & b such that $a+b=0$ & $a \neq 0$, then Y and θ are independent because $\begin{bmatrix} Y \\ \theta \end{bmatrix}$ is a bivariate Normal too.

So take $a = 1$, $b = -1 \Rightarrow Y = X - \theta$ is independent of θ.

Note that $E(X - \theta) = 0$ and

$V(X-\theta) = V(X) + V(\theta) - 2\,\mathrm{cov}(X, \theta) = 2 + 1 - 2 \times 1 = 1$

$\Rightarrow X - \theta \sim N(0, 1)$.

$$\frac{X}{\theta} = \frac{X - \theta + \theta}{\theta} = \frac{X-\theta}{\theta} + 1$$

Denote $W = \frac{X-\theta}{\theta}$. Because $X - \theta \sim N(0,1)$, $\theta \sim N(0,1)$ and $X - \theta$ is independent of θ, so W has a standard Cauchy distribution.

And $\frac{X}{\theta} = c \cdot W + d$ where $c = d = 1$.

Part II:

4. Since $|\delta(u_i)| \le 1 \Rightarrow \left|\prod_{i=1}^{n} \delta(u_i)\right| \le 1$

$$\Rightarrow f(u) \ge 0.$$

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(u)\,du = \int_0^1 \cdots \int_0^1 \left[1 - \prod_{i=1}^{n} \delta(u_i)\right] du_1 \cdots du_n$$

$$= 1 - \underbrace{\int_0^1 \cdots \int_0^1}_{n-1} \left[\prod_{i=1}^{n-1} \delta(u_i) \int_0^1 \delta(u_n)\,du_n\right] du_1 \cdots du_{n-1}$$

$$= 1 - \int_0^1 \cdots \int_0^1 \prod_{i=1}^{n-1} \delta(u_i) \times 0 \; du_1 \cdots du_{n-1} = 1$$

$\Rightarrow f(u)$ is a density function.

5.

5(a) Without loss of generality, let's take the subset $(u_1, \dots, u_{n-1})$.

$$f(u_1, \dots, u_{n-1}) = \int_0^1 f(u_1, \dots, u_{n-1}, u_n)\,du_n$$

$$= \int_0^1 \left(1 - \prod_{i=1}^{n} \delta(u_i)\right) du_n = 1 - \prod_{i=1}^{n-1} \delta(u_i) \int_0^1 \delta(u_n)\,du_n$$

$$= 1$$

And
$$f(u_1) = \underbrace{\int_0^1 \cdots \int_0^1}_{n-1} f(u_1, \dots, u_n)\,du_2 \cdots du_n$$

$$= 1 - \underbrace{\int_0^1 \cdots \int_0^1}_{n-2} \left[\prod_{i=1}^{n-1} \delta(u_i) \int_0^1 \delta(u_n)\,du_n\right] du_2 \cdots du_{n-1}$$

$$= 1 - 0 = 1$$

Similarly $f(u_i) = 1$ for $0 < u_i < 1$ for all $i = 1, \dots, n-1$.

So $f(u_1, \dots, u_{n-1}) = f(u_1) \times f(u_2) \times \cdots \times f(u_{n-1})$.

I.E. $[u_1, \dots, u_{n-1}]$ are iid Uniform $(0,1)$.

5(b). Take $\delta(u) = \begin{cases} \cos \pi u & 0 < u < 1 \\ 0 & \text{o.w.} \end{cases}$

Easy to check $|\delta(u)| \le 1$, and $\int_0^1 \delta(u)\,du = 0$.

If $(U_1, U_2, \cdots U_n)$ were independent, we should have

$f(U_1, U_2, \cdots, U_n) = \prod_{i=1}^{n} f(U_i)$.

Apparently when all $0 < U_i < 1$,

$$f(U_1, \cdots, U_n) = 1 - \prod_{i=1}^{n} \cos \pi (U_j) \ne 1 = \prod_{i=1}^{n} f(U_i)$$

except that all $U_i \equiv \frac{1}{2}$.

The set $\{(U_1, \cdots, U_n) \mid U_i \ne \frac{1}{2} \text{ for some } i = 1, \cdots, n\}$ does not have Lebesgue measure $0 \Rightarrow U_1, \cdots, U_n$ are NOT independent.

6.

6(a). Let $U = [U_1, U_2, \cdots, U_n]^T$ be a random variable with the density $f(u)$, define $X_i = F_i^{-1}(U_i)$ where $F_i^{-1}(\cdot)$ is the inverse cdf of the function $F_i(\cdot)$.

The cdf function for $X_1, X_2, \cdots, X_n$ is

$$P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n) = P(F_1^{-1}(U_1) \le x_1, \cdots, F_n^{-1}(U_n) \le x_n)$$

$$= P(U_1 \le F_1(x_1), \cdots, U_n \le F_n(x_n))$$

$$= \int_{-\infty}^{F_1(x_1)} \cdots \int_{-\infty}^{F_n(x_n)} f(u)\, du$$

$$= G(F_1(x_1), \cdots, F_n(x_n)) \triangleq H(x_1, x_2, \cdots, x_n)$$

So $H(x_1, \cdots, x_n)$ is a cdf for a random variable $X = [F_1^{-1}(U_1), \cdots, F_n^{-1}(U_n)]$ where $[U_1, \cdots, U_n]^T$ is a random variable with the density $f(u)$.

$$H(x_1) = H(x_1, +\infty, \cdots, +\infty) = G(F_1(x_1), F_2(+\infty), \cdots, F_n(+\infty))$$

$$= \int_0^{F_1(x_1)} \Big[\underbrace{\int_0^1 \cdots \int_0^1}_{n-1} \Big(1 - \prod_{i=1}^{n} b(u_i)\Big) du_2 \cdots du_n \Big] du_1$$

$$= \int_0^{F_1(x_1)} 1\, du_1 = F_1(x_1)$$

Similarly can show $H(x_i) = F_i(x_i)$ for any $i = 1, \cdots, n$.

6(b).

First derive $H(x_1, x_2, \cdots, x_{n-1})$:

$$H(x_1, x_2, \cdots, x_{n-1}) = H(x_1, x_2, \cdots, x_{n-1}, +\infty)$$

$$= G(F_1(x_1), \cdots, F_{n-1}(x_{n-1}), F_n(+\infty))$$

$$= \int_0^{F_1(x_1)} \cdots \int_0^{F_{n-1}(x_{n-1})} \left[\int_0^1 \left(1 - \prod_{i=1}^{n} \delta(u_i)\right) du_n\right] du_1 \cdots du_{n-1}$$

$$= \int_0^{F_1(x_1)} \cdots \int_0^{F_{n-1}(x_{n-1})} 1 \cdot du_1 \cdots du_{n-1}$$

$$= F_1(x_1) \times \cdots \times F_{n-1}(x_{n-1})$$

$$= H(x_1) \times \cdots \times H_{n-1}(x_{n-1})$$

$\Rightarrow$ So $(X_1, \cdots, X_{n-1})$ are independent.

Part III

7. $$X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \Sigma \right)$$

where $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$.

$\Rightarrow$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \Big| X_3 \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + A\sigma_{33}^{-1}(X_3 - \mu_3),\ B - A\sigma_{33}^{-1}A^T \right)$$

where $A = \begin{pmatrix} \sigma_{13} \\ \sigma_{23} \end{pmatrix}$, $B = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$.

The covariance matrix

$$B - A\sigma_{33}^{-1}A^T$$

$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} - \begin{bmatrix} \sigma_{13} \\ \sigma_{23} \end{bmatrix} \frac{1}{\sigma_{33}} \begin{bmatrix} \sigma_{13} & \sigma_{23} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} - \frac{\sigma_{13}^2}{\sigma_{33}} & \sigma_{12} - \frac{\sigma_{13}\sigma_{32}}{\sigma_{33}} \\ \sigma_{12} - \frac{\sigma_{13}\sigma_{23}}{\sigma_{33}} & \sigma_{22} - \frac{\sigma_{23}\sigma_{32}}{\sigma_{33}} \end{bmatrix}$$

So X_1 & X_2 are conditionally independent given X_3

$$\Leftrightarrow \quad \sigma_{12} - \frac{\sigma_{13}\sigma_{32}}{\sigma_{33}} = 0$$

$$\Leftrightarrow \quad \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} = \frac{\sigma_{13}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{33}}} \frac{\sigma_{32}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{33}}}$$

$$\Leftrightarrow \quad \rho_{12} = \rho_{13}\rho_{23}.$$

Part IV:

8. Define $f(X_1) = 1_{(-\infty, x_0]}(X_1)$.

$E[f(X_1)] = P(X_1 \le x_0) = P(X_1 - \theta \le x_0 - \theta) = \Phi(x_0 - \theta)$

the last equality holds because $X_1 - \theta \sim N(0,1)$.

The likelihood function is

$$\prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2}}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^{n} x_i^2}{2}} e^{-\frac{n\theta^2}{2}} \cdot e^{-\left(\sum_{i=1}^{n} x_i\right)\theta}.$$

This is an exponential family and $\{\theta : \theta \in R^1\}$ contains an open set in R^1, so $\sum_{i=1}^{n} X_i$ (or $\overline{X}$) is a sufficient complete statistic.

$E[1_{(-\infty, x_0)}(X_1) | \overline{X}] = P(X_1 \le x_0 | \overline{X}) = P(X_1 - \overline{X} \le x_0 - \overline{X} | \overline{X})$

We know $X_1 - \overline{X}$ is a Normal, and

$$E(X_1 - \overline{X}) = 0$$

$$V(X_1 - \overline{X}) = V\left(\left(1 - \frac{1}{n}\right)X_1 - \frac{1}{n}\sum_{i=2}^{n} X_i\right) = \frac{n-1}{n}$$

$\Rightarrow$ $X_1 - \overline{X} \sim N\left(0, \frac{n-1}{n}\right)$ which is an ancillary statistic.

Then by Basu's Theorem, $\overline{X}$ is a complete sufficient statistic and $X_1 - \overline{X}$ is an ancillary statistic $\Rightarrow$ $X_1 - \overline{X}$ is independent of $\overline{X}$.

So $E[1_{(-\infty, x_0)}(X_1) | \overline{X}] = P(X_1 - \overline{X} \le x_0 - \overline{X})$

$$= P\left(\frac{X_1 - \overline{X}}{\sqrt{\frac{n-1}{n}}} \le \frac{x_0 - \overline{X}}{\sqrt{\frac{n-1}{n}}}\right) = \Phi\left(\frac{x_0 - \overline{X}}{\sqrt{\frac{n-1}{n}}}\right),$$

which is a function of $\overline{X}$ and unbiased to $\Phi(x_0 - \theta)$.

Thus $\Phi\left(\frac{x_0 - \overline{X}}{\sqrt{\frac{n-1}{n}}}\right)$ is the UMVUE for $\Phi(x_0 - \theta)$.

Part V:

9. The likelihood function is

$$f(x_1, \cdots, x_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \theta^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2\theta}}$$

$$\log L_n(\theta|x) = n \ln \sqrt{2\pi} - \frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^n (x_i-\theta)^2$$

$$\frac{\partial \log L_n(\theta|x)}{\partial\theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n (x_i-\theta)^2 + \frac{1}{\theta}\sum_{i=1}^n (x_i-\theta)$$

Setting $\frac{\partial \log L_n(\theta|x)}{\partial\theta} = 0$ gives $n\theta^2 + n\theta - \sum_{i=1}^n x_i^2 = 0$

so $\hat{\theta}_n = -\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{n}\sum_{i=1}^n x_i^2}$ or

$\hat{\theta}_n = -\frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{n}\sum_{i=1}^n x_i^2}$

10. Since the parameter space $\theta > 0$, then $\hat{\theta}_n = -\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{n}\sum_{i=1}^n x_i^2}$

By LLN,

$$\frac{1}{n}\sum_{i=1}^n x_i^2 \xrightarrow{p} E[x_i^2] = \theta^2 + \theta.$$

Define $g(x) = -\frac{1}{2} + \sqrt{\frac{1}{4} + x}$. By continuous mapping theorem,

$$g\left(\frac{1}{n}\sum_{i=1}^n x_i^2\right) = \hat{\theta}_n \xrightarrow{p} g(\theta^2+\theta) = -\frac{1}{2} + \sqrt{\frac{1}{4} + \theta + \theta^2} = \theta.$$

Thus we obtain $\hat{\theta}_n \xrightarrow{p} \theta$.

11. Since $X_1, \dots, X_n$ are iid $N(\theta, \theta)$, by the CLT,

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\theta^2+\theta)\right) \xrightarrow{d} N(0, Var(X_i^2))$$

And we know
$$g'(x) = \frac{1}{2}\frac{1}{\sqrt{x+\frac{1}{4}}}$$
$$g'(x)\Big|_{x=\theta+\theta^2} = \frac{1}{2}\frac{1}{\sqrt{\theta+\theta^2+\frac{1}{4}}} = \frac{1}{2\theta+1}$$

By Delta method,
$$\sqrt{n}\left(g\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - g(\theta^2+\theta)\right) \xrightarrow{d} N\left(0, \left(\frac{1}{2\theta+1}\right)^2 Var(X_i^2)\right)$$

i.e.
$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\left(0, \left(\frac{1}{2\theta+1}\right)^2 Var(X_i^2)\right).$$

The moment generating function of X_i is
$$M_{X_i}(t) = E[e^{tX_i}] = e^{t\theta + \frac{1}{2}\theta t^2}$$

$$E[X_i^4] = M_{X_i}^{(4)}(t)\Big|_{t=0} = \theta^4 + 6\theta^3 + 3\theta^2.$$

And $E[X_i^2] = \theta^2 + \theta$

So
$$V(X_i^2) = E[X_i^4] - (E[X_i^2])^2 = \theta^4 + 6\theta^3 + 3\theta^2 - (\theta^2+\theta)^2$$
$$= 2\theta^2 + 4\theta^3$$

Thus
$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N\left(0, \frac{1}{(2\theta+1)^2} \times (2\theta^2 + 4\theta^3)\right)$$
$$\xrightarrow{d} N\left(0, \frac{2\theta^2}{2\theta+1}\right).$$

Part I (**Questions 1-11**)

Recall: The pdf of a gamma(α, β) distribution $(\alpha, \beta > 0)$ is $g(x) = x^{\alpha-1}e^{-x/\beta}/[\Gamma(\alpha)\beta^\alpha], x > 0$, and a chi-squared distribution with ν degrees of freedom is a gamma$(\nu/2, 2)$ distribution.

Let $n > 2$ and suppose $X_1, \dots, X_n$ are iid normal $N(\mu, \sigma^2)$ variables, for a parameter space $\Theta = \{(\mu, \sigma) : \mu \in \mathbb{R}, \sigma > 0\}$ consisting of mean μ and standard deviation σ parameters. Let $\bar{X}_n = \sum_{i=1}^n X_i/n$ and $S_n = \sqrt{S_n^2}$ where $S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2/(n-1)$.

1. Provide the distribution of $(X_1, \dots, X_n)$ given order statistics $X_{(1)} < \cdots < X_{(n)}$ and explain how this result shows the vector $(X_{(1)}, \dots, X_{(n)})$ is sufficient for $(\mu, \sigma) \in \Theta$.

2. Show that the vector $(\bar{X}_n, S_n)$ is complete and sufficient for $(\mu, \sigma) \in \Theta$. State any standard results that you use.

3. If the parameter space were instead $\tilde{\Theta} = \{(\mu, \sigma) : |\mu| = 1, \sigma > 0\}$, prove that $(\bar{X}_n, S_n)$ would *not* be complete and sufficient for $(\mu, \sigma) \in \tilde{\Theta}$.

Note: For all questions in Part I except this Question 3., the parameter space is Θ.

4. Give an example of an ancillary statistic based on $X_1, \dots, X_n$, iid $N(\mu, \sigma^2)$, $(\mu, \sigma) \in \Theta$.

5. Prove that $(\bar{X}_n, S_n)$ is independent of your statistic in Question 4.

6. Let Y be an unobservable random variable with the same $N(\mu, \sigma^2)$ distribution as $X_1, \dots, X_n$ (i.e., Y is not part of the data sample and may not be independent of $X_1, \dots, X_n$). Find the UMVUE of $\theta_1 = P_{\mu,\sigma}\big((Y - \mu)^2 \le 1\big)$.

7. Find the UMVUE of $\theta_2 = \mu/\sigma = \mu/\sqrt{\sigma^2}$.

8. If $\hat{\theta}_{2,n}$ denotes the MLE of $\theta_2 = \mu/\sigma$ (based on $X_1, \dots, X_n$), prove the distributional convergence

$$\sqrt{n}[\hat{\theta}_{2,n} - \theta_2] \xrightarrow{d} N\left(0, 1 + \frac{\mu^2}{2\sigma^2}\right) \quad \text{as } n \to \infty.$$

Hint: You may use that $\sqrt{n}[n^{-1/2}(n-1)^{1/2}S_n - \sigma] \xrightarrow{d} N(0, \sigma^2/2)$ as $n \to \infty$.

9. Show that the result in Question 8. implies the MLE $\hat{\theta}_{2,n}$ of $\theta_2 = \mu/\sigma$ is consistent.

10. Derive an asymptotically pivotal quantity for estimating $\theta_2 = \mu/\sigma$. Justify your answer.

11. Using the asymptotic pivot in Question 10., provide an approximate 95% lower confidence bound for θ_2.

Part II (Questions 12-15)

Let X and Y be independent Poisson random variables with respective pmfs given by

$$f_X(x|\theta) = e^{-\theta}\frac{\theta^x}{x!},\ x = 0, 1, 2, \ldots; \qquad f_Y(y|\lambda) = e^{-\lambda}\frac{\lambda^y}{y!},\ y = 0, 1, 2, \ldots;$$

with parameters $\theta, \lambda > 0$. We wish to test the hypothesis $H_0 : \theta = \lambda$ vs. $H_1 : c\theta = \lambda$ where $c > 1$ is some given/known constant.

12. Find the conditional distribution of X given $S = X + Y$ under H_0 *and* under H_1.

13. Assuming $S = 5$, find the most powerful test of size $\alpha = 1/4$ (conditional on S) for $H_0 : \theta = \lambda$ vs. $H_1 : c\theta = \lambda$.

14. For the test in Question 13. (again conditional on $S = 5$), show that the power of the test under H_1 converges to 1 as $c \to 1$.

15. For the test in Question 13. of $H_0 : \theta = \lambda$ vs. $H_1 : c\theta = \lambda$, what is the size of the test when not conditioning on S?

Part III (Questions 16-20)

Let random variables $X_1, \ldots, X_n$ have joint pdf $f(x_1, \ldots, x_n|\theta)$ depending on a real-valued parameter $\theta \in \Theta$. Let $R_T(\theta)$ denote the risk of an arbitrary estimator $T = h(X_1, \ldots, X_n)$ of θ with respect to a loss function $L(t, \theta) \geq 0$, $t \in \mathbb{R}, \theta \in \Theta$. Answer the following questions with respect to this risk and a prior pdf $\pi(\theta)$ on Θ.

16. Define the risk $R_T(\theta)$ of an estimator T of θ.

17. Define the Bayes risk of an estimator T of θ.

18. Define the Bayes estimator, say T^*, of θ.

19. If $\tilde{T}$ denotes the mini-max estimator of θ and T denotes any other estimator, then which of the following is true?

$\min_{\theta\in\Theta} R_{\tilde{T}}(\theta) \leq \min_{\theta\in\Theta} R_T(\theta)$, $\quad \min_{\theta\in\Theta} R_{\tilde{T}}(\theta) \geq \min_{\theta\in\Theta} R_T(\theta)$, $\quad \min_{\theta\in\Theta} R_{\tilde{T}}(\theta) \leq \max_{\theta\in\Theta} R_T(\theta)$,

$\max_{\theta\in\Theta} R_{\tilde{T}}(\theta) \leq \max_{\theta\in\Theta} R_T(\theta)$, $\quad \max_{\theta\in\Theta} R_{\tilde{T}}(\theta) \geq \max_{\theta\in\Theta} R_T(\theta)$, $\quad \min_{\theta\in\Theta} R_{\tilde{T}}(\theta) \leq \max_{\theta\in\Theta} R_T(\theta)$.

20. If an estimator T^* is the Bayes estimator and has constant risk (i.e., $R_{T^*}(\theta) = c > 0$ for any $\theta \in \Theta$), then show T^* must be the mini-max estimator of θ.

Part IV (**Questions 21-24**)

Let $X_1, \ldots, X_n$ be iid Geometric(p) random variables with pmf $f_X(x|p) = p(1-p)^{x-1}$, $x = 1, 2, 3, \ldots$ for a parameter $0 < p < 1$. It holds that mean $E_p(X_1) = 1/p$ and variance $\text{Var}_p(X_1) = (1-p)/p^2$.

21. Show that the method of moments (MOM) estimator $\hat{\theta}_n$ of $\theta = p(1-p)$ is asymptotically normal (with proper centering and scaling).

22. Show that the estimator $\tilde{\theta}_n = \sum_{i=1}^n \mathbb{I}(X_i = 2)/n$ of θ is also asymptotically normal (with proper centering and scaling), where $\mathbb{I}(\cdot)$ denotes the indicator function.

23. The UMVUE of θ (based on $X_1, \ldots, X_n$) is

$$T_n = \frac{(n-1)(S_n - n)}{(S_n - 1)(S_n - 2)},$$

where $S_n = \sum_{i=1}^n X_i$. Prove that $\sqrt{n}(T_n - \hat{\theta}_n) \xrightarrow{p} 0$ as $n \to \infty$, where $\hat{\theta}_n$ is the MOM estimator of θ.

24. In comparing limiting variances, prove that the estimators T_n and $\hat{\theta}_n$ are asymptotically more efficient than $\tilde{\theta}_n$.

Hint: The limiting distribution of T_n is needed for this.

1. Conditional on the order statistics $(X_{(1)}, \ldots, X_{(n)}) = (x_{(1)}, \ldots, x_{(n)})$, the potential values of $(X_1, \ldots, X_n)$ correspond to the set $\mathcal{S}$ of $\mathbb{R}^n$-valued vectors consisting of all possible permutations of $(x_{(1)}, \ldots, x_{(n)})$ where, by the iid property of $X_1, \ldots, X_n$, any permutation is equally likely. Hence, the (discrete) conditional distribution of $(X_1, \ldots, X_n)$ given $(X_{(1)}, \ldots, X_{(n)}) = (x_{(1)}, \ldots, x_{(n)})$ is

$$P_{\mu,\sigma}(X_1 = x_1, \ldots, X_n = x_n | x_{(1)}, \ldots, x_{(n)}) = \begin{cases} \dfrac{1}{n!} & \text{if } (x_1, \ldots, x_n) \in \mathcal{S} \\ 0 & \text{othewise,} \end{cases}$$

which is free of parameters (μ, σ). Hence, by definition, $(X_{(1)}, \ldots, X_{(n)})$ is sufficient for $(\mu, \sigma) \in \Theta$.

2. The joint distribution of $(X_1, \ldots, X_n)$ has an exponential family form: for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$, it holds that

$$\begin{aligned} f(x_1, \ldots, x_n | \mu, \sigma) &= \prod_{i=1}^n f(x_i | \mu, \sigma) \\ &= \prod_{i=1}^n (2\pi)^{-1/2} \sigma^{-1} \exp\left[-(X_i - \mu)^2/(2\sigma^2)\right] \\ &= h(x_1, \ldots, x_n) C(\mu, \sigma) \exp\left[q_1(\mu, \sigma) \sum_{i=1}^n x_i^2 + q_2(\mu, \sigma) \sum_{i=1}^n x_i\right] \end{aligned}$$

where $h(x_1, \ldots, x_n) = 1$,

$$C(\mu, \sigma) = (2\pi)^{-n/2} \sigma^{-n} \exp[-\mu^2/(2\sigma^2)], \quad q_1(\mu, \sigma) = -\frac{1}{2\sigma^2}, \quad q_2(\mu, \sigma) = \frac{\mu}{\sigma^2}.$$

The subset of $\mathbb{R}^2$ given by

$$\Big\{[q_1(\mu, \sigma), q_2(\mu, \sigma)] : (\mu, \sigma) \in \Theta\Big\} = \Big\{[-1/(2\sigma^2), \mu/\sigma^2] : \mu \in \mathbb{R}, \sigma > 0\Big\} = (-\infty, 0) \times \mathbb{R}$$

contains an open subset in $\mathbb{R}^2$. Hence, $(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i)$ is complete and sufficient for (μ, σ). Statistics $\sum_{i=1}^n X_i^2$ and $\sum_{i=1}^n X_i$ have a one-to-one correspondence with statistics $n^{-1} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2$ and $\bar{X}_n$, which have a one-to-one correspondence with $(\bar{X}_n, S_n)$. Therefore, $(\bar{X}_n, S_n)$ must be complete and sufficient for $(\mu, \sigma) \in \Theta$.

3. Consider a function (statistic) of $(\bar{X}_n, S_n)$ given by $h[(\bar{X}_n, S_n)] = \bar{X}_n^2 - 1 - n^{-1} S_n^2$. Then, for all $(\mu, \sigma) \in \tilde{\Theta}$ (where $|\mu| = 1$), it holds that

$$\begin{aligned} \mathrm{E}_{\mu,\sigma} h[(\bar{X}_n, S_n)] &= \mathrm{E}_{\mu,\sigma}(\bar{X}_n^2) - n^{-1} \mathrm{E}_{\mu,\sigma}(S_n^2) - 1 \\ &= \mathrm{Var}_{\mu,\sigma}(\bar{X}_n^2) + [\mathrm{E}_{\mu,\sigma} \bar{X}_n]^2 - n^{-1}\sigma^2 - 1 \\ &= n^{-1}\sigma^2 + \mu^2 - n^{-1}\sigma^2 - 1 = 0, \end{aligned}$$

using that $\bar{X}_n$ and S_n^2 are unbiased for μ and σ^2, respectively. While $\mathrm{E}_{\mu,\sigma} h[(\bar{X}_n, S_n)] = 0$ then holds for all $(\mu, \sigma) \in \tilde{\Theta}$, this does not in turn imply $P_{\mu,\sigma}(h[(\bar{X}_n, S_n)] = 0) = 1$ for all $(\mu, \sigma) \in \tilde{\Theta}$. In fact, $P_{\mu,\sigma}(h[(\bar{X}_n, S_n)] = 0) = 0$ for all $(\mu, \sigma) \in \tilde{\Theta}$. Hence, $(\bar{X}_n, S_n)$ would not be complete (but is sufficient) for $(\mu, \sigma) \in \tilde{\Theta}$.

4. No matter the data-generating parameters $(\mu, \sigma) \in \Theta$, the variables $Z_i = (X_i - \mu)/\sigma$, $i = 1, \ldots, n$ are iid $N(0,1)$ variables so that, for example, the statistic $(X_1 - X_2)/(X_1 - X_3) = (Z_1 - Z_2)/(Z_1 - Z_3)$ has a distribution free of (μ, σ) and is therefore an ancillary statistic. There are plenty of other examples of ancillary statistics, e.g., $(\bar{X}_n - X_n)/S_n$.

5. As $(\bar{X}_n, S_n)$ is complete and sufficient for $(\mu, \sigma) \in \Theta$, any ancillary statistic will be independent of $(\bar{X}_n, S_n)$ (regardless of the data-generating parameters (μ, σ)) by Basu's theorem.

6. $Y - \mu \sim N(0, \sigma^2)$ has the same distribution as $(X_1 - X_2)/\sqrt{2}$, so that $T = \mathbb{I}[(X_1 - X_2)^2 \leq 2]$ is an unbiased estimator of θ_1, i.e.,
$$\theta_1 = P_{\mu,\sigma}[(Y-\mu)^2 \leq 1] = P_{\mu,\sigma}\left[((X_1 - X_2)^2/2 \leq 1\right] = \mathrm{E}_{\mu,\sigma} T.$$
As $(\bar{X}_n, S_n)$ is complete/sufficient, the UMVUE of θ_1 can be then found as
$$\begin{aligned} \mathrm{E}[T|\bar{X}_n, S_n] &= P\left((X_1 - X_2)^2 \leq 2 | \bar{X}_n, S_n\right) \\ &= P\left(\frac{(X_1 - X_2)^2}{S_n^2} \leq \frac{2}{S_n^2}\Big|\bar{X}_n, S_n\right) \\ &= P\left(V \leq \frac{2}{S_n^2}\Big|\bar{X}_n, S_n\right), \end{aligned}$$
where $V = (X_1 - X_2)^2/S_n^2 = (Z_1 - Z_2)(n-1)/\sum_{i=1}^n (Z_i - \bar{Z}_n)^2$ is ancillary and independent of $\bar{X}_n, S_n$ (here $Z_i = (X_i - \mu)/\sigma$). If $G(\cdot)$ denotes the marginal cdf of V, then (by independence) this conditional probability is equal to the marginal probability that $V \leq 2/S_n^2$ when treating S_n as fixed/given, or
$$\mathrm{E}[T|\bar{X}_n, S_n] = G\left(\frac{2}{S_n^2}\right);$$
this is the UMVUE of θ_1. Answers may vary in expressing the UMVUE, but the UMVUE is unique.

7. By normal theory, $\bar{X}_n$ and S_n^2 are independent and $W = (n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$. Hence,
$$\begin{aligned} \mathrm{E}_{\mu,\sigma} \frac{\bar{X}_n}{S_n} &= \mathrm{E}_{\mu,\sigma}(\bar{X}_n) \cdot \mathrm{E}_{\mu,\sigma}\left(\frac{1}{S_n}\right) \\ &= \mu \cdot \mathrm{E}_{\mu,\sigma}\left(\frac{(n-1)^{1/2}}{\sigma W^{1/2}}\right) \\ &= \frac{\mu(n-1)^{1/2}}{\sigma} \int_0^\infty \frac{1}{\Gamma[(n-1)/2]2^{(n-1)/2}} w^{-1/2} w^{\frac{n-1}{2}-1} e^{-w/2} dw \\ &= \frac{\mu(n-1)^{1/2}}{\sigma} \frac{\Gamma[(n-2)/2]2^{(n-2)/2}}{\Gamma[(n-1)/2]2^{(n-1)/2}} \end{aligned}$$
Hence,
$$T_1 = \frac{2^{1/2}}{(n-1)^{1/2}} \frac{\Gamma[(n-1)/2]}{\Gamma[(n-2)/2]} \frac{\bar{X}_n}{S_n}$$
is a function of the complete/sufficient $(\bar{X}_n, S_n)$ and unbiased for $\theta_2 = \mu/\sigma$; therefore, T_1 is the UMVUE of θ_2.

8. The well-known MLE of (μ, σ) is $(\bar{X}_n, \tilde{S}_n \equiv S_n n^{-1/2}(n-1)^{1/2})$ so that the MLE of $\theta_2 = \mu/\sigma$ is $\hat{\theta}_{2,n} = \bar{X}_n/\tilde{S}_n$. Write

$$\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) = \sqrt{n}(\hat{\theta}_{2,n} \pm \mu/\tilde{S}_n - \theta_2) = \sqrt{n}\frac{(\bar{X}_n - \mu)}{\tilde{S}_n} + \sqrt{n}\mu(\tilde{S}_n^{-1} - \sigma^{-1}) = I_{1,n} + I_{2,n}.$$

As $\tilde{S}_n \xrightarrow{p} \sigma$ and $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$, by Slutsky's theorem, we have $I_{1,n} \xrightarrow{d} L_1 \sim N(0,1)$. Using the delta method with $\sqrt{n}(\tilde{S}_n - \sigma) \xrightarrow{d} N(0, \sigma^2/2)$, we have $\sqrt{n}(\tilde{S}_n^{-1} - \sigma^{-1}) = \sqrt{n}[g(\tilde{S}_n) - g(\sigma)] \xrightarrow{d} N(0, [g'(\sigma)]^2\sigma^2/2 = \sigma^{-2}/2)$, using $g(x) = x^{-1}$ and $g'(x) = -x^{-2}$. Hence, $I_{2,n} \xrightarrow{d} L_2 \sim N(0, \mu^2\sigma^{-2}/2)$. As $I_{1,n}$ and $I_{2,n}$ are independent by normal theory, so are their limits L_1 and L_2 and so $\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \xrightarrow{d} L_1 + L_2 \sim N(0, 1 + \mu^2\sigma^{-2}/2)$.

There are other ways to solve this. For example, using the Delta method, one can Taylor expand the estimator $\hat{\theta}_{2,n} = \bar{X}_n/\tilde{S}_n = h(\bar{X}_n, \tilde{S}_n)$ (for $h(x,y) = x/y$) around (μ, σ) as

$$\hat{\theta}_{2,n} = h(\bar{X}_n, \tilde{S}_n) \approx h(\mu, \sigma) + \frac{\partial h(\mu, \sigma)}{\partial \mu}(\bar{X}_n - \mu) + \frac{\partial h(\mu, \sigma)}{\partial \sigma}(\tilde{S}_n - \sigma)$$

so that, using $h(\mu, \sigma) = \theta_2$,

$$\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \approx \frac{1}{\sigma}\sqrt{n}(\bar{X}_n - \mu) - \frac{\mu}{\sigma^2}\sqrt{n}(\tilde{S}_n - \sigma) \xrightarrow{d} L_1 + L_2.$$

To say $\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \approx \frac{1}{\sigma}\sqrt{n}(\bar{X}_n - \mu) - \frac{\mu}{\sigma^2}\sqrt{n}(\tilde{S}_n - \sigma)$, technically means that the remainder

$$R_n = \sqrt{n}(\hat{\theta}_{2,n} - \theta_2) - \frac{1}{\sigma}\sqrt{n}(\bar{X}_n - \mu) + \frac{\mu}{\sigma^2}\sqrt{n}(\tilde{S}_n - \sigma)$$

converges to zero in probability.

9. By Slutsky's theorem (with $n^{-1/2} \to 0$) and the result in Question 8.,

$$(\hat{\theta}_{2,n} - \theta_2) = \frac{1}{\sqrt{n}}\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \xrightarrow{d} 0 \cdot N(0, 1 + \mu^2\sigma^{-2}/2) = 0$$

so that $(\hat{\theta}_{2,n} - \theta_2) \xrightarrow{d} 0$ or, equivalently, $(\hat{\theta}_{2,n} - \theta_2) \xrightarrow{p} 0$. Hence, $\hat{\theta}_{2,n} \xrightarrow{p} \theta_2$.

10. By Question 8., $\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \xrightarrow{d} N(0, 1 + \mu^2\sigma^{-2}/2)$ with limiting variance $1 + \mu^2\sigma^{-2}/2 = 1 + \theta_2^2/2$. By Question 9., $\hat{\theta}_{2,n} \xrightarrow{p} \theta_2$ so that, as $g(x) = 1 + x^2/2$ is a continuous function of $x \in \mathbb{R}$, we have $g(\hat{\theta}_{2,n}) = 1 + \hat{\theta}_{2,n}^2/2 \xrightarrow{p} g(\theta_2) = 1 + \theta_2^2/2$ by the continuous mapping theorem. We then have

$$\frac{1}{\sqrt{1 + \hat{\theta}_{2,n}^2/2}}\sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \xrightarrow{d} N(0,1)$$

by Slutsky's theorem; $\sqrt{n}(\hat{\theta}_{2,n} - \theta_2)/\sqrt{1 + \hat{\theta}_{2,n}^2/2}$ is asymptotically pivotal.

11. By the solution to Question 10., for a 95% lower bound for θ_2 with large n, use (for $Z \sim N(0,1)$, $z_{0.95} = 1.96$)

$$\begin{aligned} 0.95 &= P(Z \le z_{0.95}) \\ &\approx P\left(\frac{1}{\sqrt{1+\hat{\theta}_{2,n}^2/2}} \sqrt{n}(\hat{\theta}_{2,n} - \theta_2) \le z_{.95} \right) \\ &= P\left(\hat{\theta}_{2,n} - \frac{z_{0.95}}{\sqrt{n}} \sqrt{1+\hat{\theta}_{2,n}^2/2} \le \theta_2 \right). \end{aligned}$$

That is, an approximate 95% lower bound for θ_2 is $\hat{\theta}_{2,n} - z_{0.95}\frac{1}{\sqrt{n}}\sqrt{1+\hat{\theta}_{2,n}^2/2}$.

12. By the independence of X and Y, $S = X + Y$ is Poisson$(\theta + \lambda)$ distributed with pmf

$$f_S(s|\theta+\lambda) = e^{-\theta-\lambda}\frac{(\theta+\lambda)^s}{s!}, \quad s = 0, 1, 2, \ldots.$$

Given $S = s \in \{0, 1, 2, \ldots\}$, the conditional distribution of X is

$$\begin{aligned} P_{\theta,\lambda}(X = x|S = s) &= \frac{P_{\theta,\lambda}(X = x, S = s)}{f_S(s|\theta+\lambda)} \\ &= \frac{f_X(x|\theta) f_Y(s-x|\lambda)}{f_S(s|\theta+\lambda)} \\ &= \frac{e^{-\theta}\dfrac{\theta^x}{x!} e^{-\lambda}\dfrac{\lambda^{(s-x)}}{(s-x)!}}{e^{-\theta-\lambda}\dfrac{(\theta+\lambda)^s}{s!}} \\ &= \binom{s}{x}(p)^x(1-p)^{s-x} \end{aligned}$$

for $x = 0, 1, \ldots, s$ and $p = \theta/(\theta+\lambda)$; for any other x value, $P_{\theta,\lambda}(X = x|S = s) = 0$.

Hence, under $H_0 : \theta = \lambda$, it follows that $X|S = s \sim$Binomial$(s, 1/2)$ while, under $H_1 : c\theta = \lambda$, we have $X|S = s \sim$Binomial$(s, [1+c]^{-1})$.

13. The most powerful (MP) test (conditional on S) of $H_0 : \theta = \lambda$ vs $H_1 : c\theta = \lambda$ is the MP test of $H_0 : p = 1/2$ vs $H_1 : p = 1/(c+1)$ based on the conditional pmf of X given $S = 5$, which is

$$f(x|p) = \binom{5}{x} p^x (1-p)^x, \quad x = 0, 1, \ldots, 5.$$

From the Neyman-Pearson lemma, the MP test given $S = 5$ of size $\alpha = 1/4$ has the form

$$\phi(X) = \begin{cases} 1 & \text{if } f(X|[1+c]^{-1}) > kf(X|1/2) \\ \gamma & \text{if } f(X|[1+c]^{-1}) = kf(X|1/2) \\ 0 & \text{if } f(X|[1+c]^{-1}) < kf(X|1/2) \end{cases}$$

for some $k > 0$ and $\gamma \in [0,1]$ such that $\mathrm{E}_{X|S=5,p=1/2}[\phi(X)] = 1/4$. Note for $k > 0$,

$$f(x|[1+c]^{-1}) \overset{>}{\underset{<}{=}} kf(x|1/2) \Leftrightarrow \left(\frac{1}{1+c}\right)^x \left(\frac{c}{1+c}\right)^{5-x} \overset{>}{\underset{<}{=}} k\left(\frac{1}{2}\right)^5$$

$$\Leftrightarrow \left(\frac{1}{c}\right)^x \overset{>}{\underset{<}{=}} k\left(\frac{c+1}{2c}\right)^5 \equiv k_1 \Leftrightarrow -x\log c \overset{>}{\underset{<}{=}} \log k_1 \equiv k_2 \Leftrightarrow x \overset{<}{\underset{>}{=}} k_2/[-\log c] \equiv k_3$$

Hence, the MP test has the form

$$\phi(X) = \begin{cases} 1 & \text{if } X < k_3 \\ \gamma & \text{if } X = k_3 \\ 0 & \text{if } X > k_3 \end{cases}$$

for some k_3. Under H_0, the conditional pdf of X given $S = 5$ has values $\frac{1}{32}, \frac{5}{32}, \frac{10}{32}, \frac{10}{32}, \frac{10}{32}, \frac{1}{32}$ for $x = 0, 1\ldots, 5$, so pick $k_3 = 2$ so that

$$1/4 = \mathrm{E}_{X|S=5,p=1/2}[\phi(X)] = P(X < 2|S = 5, 1/2) + \gamma P(X = 2|S = 5, 1/2) = \frac{3}{16} + \gamma\frac{5}{16},$$

implying $\gamma = 1/5$.

14. Under H_1, the power function (conditional on $S = 5$) satisfies

$$\begin{aligned} & \mathrm{E}_{X|S=5,p=[1+c]^{-1}}[\phi(X)] \\ = \; & P(X < 2|S = 5, [1+c]^{-1}) + \gamma P(X = 2|S = 5, [1+c]^{-1}) \\ = \; & \left(\frac{c}{c+1}\right)^5 + 5\left(\frac{1}{c+1}\right)\left(\frac{c}{c+1}\right)^4 + \frac{1}{5}10\left(\frac{1}{c+1}\right)\left(\frac{c}{c+1}\right)^3 \\ = \; & \frac{c^5 + 5c^4 + 2c^3}{(c+1)^5} \\ = \; & \frac{1 + 5c^{-1} + 2c^{-2}}{(c^{-1}+1)^5} \\ \to \; & \frac{1 + 5\cdot 0 + 2\cdot 0}{(0+1)^5} = 1 \end{aligned}$$

as $c \to \infty$.

15. Unconditionally, the test
$$\phi(X) = \begin{cases} 1 & \text{if } X < 2 \\ 1/5 & \text{if } X = 2 \\ 0 & \text{if } X > 2 \end{cases}$$
has an expected value, depending $X \sim$Poisson(θ) for $\theta > 0$, given by
$$\mathrm{E}_\theta \phi(X) = e^{-\theta} + e^{-\theta}\theta + e^{-\theta}\theta^2/10.$$
Note any value of $\theta > 0$ is allowed under $H_0 : \theta = \lambda$ so that the size of the test is
$$\sup_{\theta > 0}[e^{-\theta} + e^{-\theta}\theta + e^{-\theta}\theta^2/10] = 1,$$
which is found by $\lim_{\theta \to 0}[e^{-\theta} + e^{-\theta}\theta + e^{-\theta}\theta^2/10] = 1$.

16. Risk is expected loss under the data-generating parameter θ:
$$R_T(\theta) = \mathrm{E}_\theta L(T, \theta) = \int_{\mathbb{R}^n} L(h(X_1, \ldots, X_n), \theta) f(x_1, \ldots, x_n|\theta) dx_1 \ldots dx_n$$

17. Bayes risk, BR_T, is prior-weighed risk:
$$BR_T = \int_\Theta R_T(\theta)\pi(\theta)d\theta$$

18. The Bayes estimator T^* has minimal Bayes risk compared to any other estimator T:
$$BR_{T^*} = \min_{\text{any } T} BR_T$$

19. The mini-max estimator $\tilde{T}$ has no greater maximum risk compared to any other estimator T:
$$\max_{\theta \in \Theta} R_{\tilde{T}}(\theta) \le \max_{\theta \in \Theta} R_T(\theta)$$

20. Pick any other estimator T. If T^* is the Bayes estimator then $BR_{T^*} \le BR_T$ and if $R_{T^*}(\theta) = c$ for all $\theta \in \Theta$, then
$$\begin{aligned} \max_{\theta \in \Theta} R_{T^*}(\theta) = c = c\int_\Theta \pi(\theta)d\theta &= \int_\Theta R_{T^*}(\theta)\pi(\theta)d\theta \\ &= BR_{T^*} \\ &\le BR_T \\ &= \int_\Theta R_{T^*}(\theta)\pi(\theta)d\theta \\ &\le \max_{\theta \in \Theta} R_T(\theta) \int_\Theta \pi(\theta)d\theta \\ &= \max_{\theta \in \Theta} R_T(\theta). \end{aligned}$$

Hence, T^* must be the mini-max estimator.

21. The MOM estimator $\hat{p}_n$ of p solves $1/\hat{p}_n = \mathrm{E}_{\hat{p}_n}(X_1) = \bar{X}_n$ or $\hat{p}_n = 1/\bar{X}_n$. Then, the MOM estimator of $\theta = p(1-p)$ is, by definition, $\hat{\theta}_n = \hat{p}_n(1-\hat{p}_n) = \bar{X}_n^{-1}(1-\bar{X}_n^{-1})$. By the CLT, $\sqrt{n}(\bar{X}_n - p^{-1}) \xrightarrow{d} N(0, \mathrm{Var}_p(X_1) = p^{-2}(1-p))$ as $n \to \infty$. By the Delta Method for $g(x) = x^{-1}(1-x^{-1})$,
$$\sqrt{n}(\hat{\theta}_n - \theta) = \sqrt{n}[g(\bar{X}_n) - g(p^{-1})] \xrightarrow{d} N[0, [g'(p^{-1})]^2 p^{-2}(1-p) = p^2(1-p)(2p-1)^2]$$
using $g'(x) = -x^{-2}(1-x^{-1}) + x^{-3}$ so $g'(p^{-1}) = p^3 - p^2(1-p) = p^2(2p-1)$.

22. For $Y_i = \mathbb{I}(X_i = 2)$, $Y_1, \ldots, Y_n$ are iid Bernoulli(θ) as $\mathrm{E}_p Y_1 = P_p(X_1 = 2) = p(1-p) = \theta$. By the CLT,
$$\sqrt{n}(\tilde{\theta}_n - \theta) = \sqrt{n}[\bar{Y}_n - \mathrm{E}_p Y_1] \xrightarrow{d} N[0, \mathrm{Var}_p(Y_1) = \theta(1-\theta) = p(1-p) - p^2(1-p)^2].$$

23. The UMVUE of θ can be written as
$$T_n = \frac{(n-1)(S_n - n)}{(S_n-1)(S_n-2)} = \frac{(1-n^{-1})(\bar{X}_n - 1)}{(\bar{X}_n - n^{-1})(\bar{X}_n - n^{-1}2)}$$
so that
$$\begin{aligned} T_n - \hat{\theta}_n &= \frac{(1-n^{-1})(\bar{X}_n - 1)}{(\bar{X}_n - n^{-1})(\bar{X}_n - n^{-1}2)} - \frac{1}{\bar{X}_n}\frac{\bar{X}_n - 1}{\bar{X}_n} \\ &= \frac{\bar{X}_n - 1}{\bar{X}_n^2(\bar{X}_n - n^{-1})(\bar{X}_n - n^{-1}2)} A_n, \end{aligned}$$
$$A_n = (1-n^{-1})\bar{X}_n{}^2 - (\bar{X}_n - n^{-1})(\bar{X}_n - n^{-1}2) = -n^{-1}\bar{X}_n{}^2 + 3n^{-1}\bar{X}_n - 2n^{-2}.$$
By WLLN, $\bar{X}_n \xrightarrow{p} \mathrm{E}_p(X_1) = p^{-1}$, so that by the continuous mapping theorem $\sqrt{n}A_n = n^{-1/2}(\bar{X}_n{}^2 + 3n^{-1/2}\bar{X}_n - 2n^{-3/2}) \xrightarrow{p} 0(p^{-2}+0+0) = 0$ and
$$\sqrt{n}(T_n - \hat{\theta}_n) \xrightarrow{p} \frac{p^{-1}-1}{p^{-4}} 0 = 0.$$

24. As $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N[0, p^2(1-p)(2p-1)^2]$ by Question 21. and $\sqrt{n}(T_n - \hat{\theta}_n) \xrightarrow{p} 0$ by Question 23., it follows that
$$\sqrt{n}(T_n - \theta) = \sqrt{n}(\hat{\theta}_n - \theta) - \sqrt{n}(T_n - \hat{\theta}_n) \xrightarrow{d} 0 + N[0, p^2(1-p)(2p-1)^2]$$
by Slutsky's theorem. So, T_n and $\hat{\theta}_n$ have the same limiting variance and the same efficiency. To show that T_n and $\hat{\theta}_n$ are asymptotically more efficient that $\tilde{\theta}_n$, we compare the respective limiting variances for $p \in (0,1)$,
$$\begin{aligned} p^2(1-p)(2p-1)^2 &< p(1-p) - p^2(1-p)^2 \\ &\Leftrightarrow \\ p(2p-1)^2 &< 1 - p(1-p) \\ &\Leftrightarrow \\ 0 &< 1 - 2p + 5p^2 - 4p^3 \equiv h(p). \end{aligned}$$
Taking a derivative $h'(p) = -2 + 10p - 12p^2 = -2(1-5p+6p^2) = -2(1-2p)(1-3p)$, so that $h(p)$ is strictly decreasing on $[0,1/3]$ and $[1/2,1]$ and strictly increasing on $[1/3,1/2]$. So the unique minimum of $h(p)$ on $[0,1]$ is $\min\{h(1/3), h(1)\} = h(1) = 0$, occurring at 1. This implies $h(p) > 0$ holds for $p \in (0,1)$. Hence, T_n and $\hat{\theta}_n$ are asymptotically more efficient.

In the following, $I\!R$ denote the set of real numbers, $I\!N$ is the set of natural numbers $\{1, 2, \ldots\}$. For any set $\mathcal{X}$ and $\mathcal{P}(\mathcal{X})$ denotes the power set of subsets of $\mathcal{X}$.

1. **(a)** Let $\Omega \neq \emptyset$. Define what it means for a collection of subsets of Ω to be:

   (i) semi-algebra, (ii) algebra, (iii) σ-algebra, (vi) π-system, (v) λ-system

   (b) A nonempty collection $\mathcal{M}$ is said to be a *monotone class* if it is closed under increasing unions and decreasing intersections (i.e. If $\{A_n\}$ is an increasing sequence of sets in $\mathcal{M}$ then $\cup_{n\geq 1} A_n \in \mathcal{M}$ and if $\{A_n\}$ is an decreasing sequence of sets in $\mathcal{M}$ then $\cap_{n\geq 1} A_n \in \mathcal{M}$). Prove that if $\mathcal{F}$ is a σ-algebra, it is also a/an

   (i) algebra, (ii) π-system, (iii) λ-system, (iv) semi-algebra, (v) monotone class.

2. Let $\{\mu_n\}$ be measures on a measurable space $(\Omega, \mathcal{F})$ and assume that $\mu_{n+1}(A) \geq \mu_n(A)$ for all $A \in \mathcal{F}, n \geq 1$. Show that $\mu \doteq \lim_{n\to\infty} \mu_n$ is a measure on $(\Omega, \mathcal{F})$.

3. Let $\mathcal{F}$ be a σ-algebra of subsets of $\Omega \neq \emptyset$. Take $\Theta \in \mathcal{P}(\Omega)$ such that $\Theta \notin \mathcal{F}$. Define $\mathcal{F}_\Theta = \{F \cap \Theta : F \in \mathcal{F}\}$. Show that $\mathcal{F}_\Theta$ is a σ-algebra of subsets of Θ.

4. Let $(I\!N, \mathcal{P}(I\!N), \mu)$ be a measure space, $r \in I\!R$ $(r \geq 1)$ and $g : I\!N \to I\!R$ is a function.
   (a) If μ is the counting measure and $g(n) = n^{\frac{-1}{p}}$ for $n \in I\!N$ (for some real $p \in [1, \infty)$), then show that $g \in L^r$ iff $p < r$.
   (b) If the measure μ is defined by $\mu(\{n\}) = \frac{1}{n^2}$ for $n \in I\!N$ and $g(n) = n^{\frac{1}{p}}$ for $n \in I\!N$ (for some real $p \in (1, \infty)$), then show that $g \in L^r$ iff $1 \leq r < p$.

5. Let $X_1, X_2, \ldots,$ be a sequence of iid random variables defined on some probability space $(\Omega, \mathcal{F}, P)$ and $S_n = \sum_{i=1}^n X_i$ for $n \geq 1$.

   (a) If $E|X_1|^2 < \infty$ and $E(X_1) = 0$, then -

   (i) State the *strong* law of large numbers for $\{X_n\}$.

   (ii) Argue, using (i), that
   $$\frac{S_n}{n} \xrightarrow{p} 0.$$

   (iii) Show the above convergence holds by showing convergence of the corresponding characteristic functions.

   (iv) If $T_n = \sum_{i=1}^n X_i^2$, $n \geq 1$ show that
   $$\frac{S_n}{\sqrt{T_n}} \xrightarrow{d} \mathcal{N}(0, 1).$$
   (Here $\mathcal{N}(0, 1)$ is the Gaussian distribution with mean zero and variance 1)

   (b) In general,

   (i) Show that $E|X_1| < \infty$ iff $\sup_{n\geq 1} \frac{|X_n|}{n} < \infty$ a.s.

   (ii) Show that $E|X_1| < \infty$ iff $\sup_{n\geq 1} \frac{|S_n|}{n} < \infty$ a.s.

1. (**a.**) (i) $\mathcal{C} \subset \mathcal{P}(\Omega)$ is semi-algebra if

- $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$;
- $\forall A \in \mathcal{C} \Rightarrow A^C = \bigcup_{i=1}^k B_i, \{B_i\}_{i=1}^k \subset \mathcal{C}$ are disjoint;
- $\emptyset \in \mathcal{P}(\Omega)$.

(ii) $\mathcal{F} \subset \mathcal{P}(\Omega)$ is an algebra if

- $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$;
- $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$;
- $\Omega \in \mathcal{F}$.

(iii) $\mathcal{F} \subset \mathcal{P}(\Omega)$ is an σ-algebra if

- $\{A_i\} \subset \mathcal{F} \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{F}$;
- $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$;
- $\Omega \in \mathcal{F}$.

(iv) $\mathcal{C} \subset \mathcal{P}(\Omega)$ is a π-system if $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$.

(v) $\mathcal{L} \subset \mathcal{P}(\Omega)$ is a λ-system if

- $\Omega \in \mathcal{L}$;
- $A, B \in \mathcal{L}, A \subset B \Rightarrow B \backslash A \in \mathcal{L}$;
- $A_n \in \mathcal{L}, A_n \subset A_{n+1} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{L}$.

(**b.**) (i) Note that being closed under countable unions implies being closed under finite unions (if $A, B \in \mathcal{F}$. then $A \cup B = \bigcup_{i=1}^\infty A_i \in \mathcal{F}$ where $A_1 = A, A_2 = B, A_i = \emptyset \in \mathcal{F}, i \geq 3$.) Hence, $\mathcal{F}$ is an algebra.

(ii) $A, B \in \mathcal{F} \Rightarrow A^C, B^C \in \mathcal{F} \Rightarrow (A \cap B)^C = A^C \cup B^C \in \mathcal{F} \Rightarrow A \cap B = ((A \cap B)^C)^C \in \mathcal{F}$. So $\mathcal{F}$ is a π-system.

(iii)

- $\mathcal{F}$ is σ-algebra $\Rightarrow \Omega \in \mathcal{F}$.
- $A, B \in \mathcal{F}, A \subset B \Rightarrow A^C \in \mathcal{F} \Rightarrow B \backslash A = B \cap A^C \in \mathcal{F}$ from the result of (ii) in (b).
- $A_n \in \mathcal{F}, A_n \subset A_{n+1} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{F}$, since sigma algebra is closed under any countable unions.

So $\mathcal{F}$ is a λ-system.

(iv)

- $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$ follows from (ii) of part (b).
- $\forall A \in \mathcal{F} \Rightarrow A^C = \bigcup_{i=1}^1 B_i, B_1 = A^C \in \mathcal{F}$;
- $\Omega \in \mathcal{F} \Rightarrow \emptyset = \Omega^C \in \mathcal{F}$.

So $\mathcal{F}$ is a semi-algebra.

(v) Sigma algebra is closed under countable unions and complements. Hence by De Morgan's law, it is closed under countable intersections.

2. Since $\mu_n(\emptyset) = 0$ for all $n \geq 1$, we have that $\mu(\emptyset) = 0$.
So we need show that $\mu(\cup_{k\geq 1} A_k) = \sum_{k\geq 1} A_k$ for any disjoint collection $\{A_k\}$. Since $\mu_n(A_k) \leq \mu(A_k)$ we have $\mu_n(\cup_{k\geq 1} A_k) = \sum_{k\geq 1} \mu_n(A_k) \leq \sum_{k\geq 1} \mu(A_k)$ for all $n \geq 1$. Hence, taking limits, we get

$$\mu(\cup_{k\geq 1} A_k) \leq \sum_{k\geq 1} \mu(A_k).$$

To get the other inequality, note that $\mu_n(\cup_{k\geq 1} A_k) = \sum_{k\geq 1} \mu_n(A_k) \geq \sum_{k\geq 1}^{m} \mu_n(A_k)$ for all integer $m \geq 1$. Hence taking limit as $n \to \infty$ (note the finite sum on the right hand side), we get $\mu(\cup_{k\geq 1} A_k) \geq \sum_{k\geq 1}^{m} \mu(A_k)$ - and this holds for all $m \geq 1$. Hence, we have

$$\mu(\cup_{k\geq 1} A_k) \geq \sum_{k\geq 1} \mu(A_k).$$

This proves $\mu(\cup_{k\geq 1} A_k) = \sum_{k\geq 1} A_k$ and that completes the proof.

3. Taking $F = \Omega \in \mathcal{F}$, we get $\Theta = \Omega \cap \Theta \in \mathcal{F}_\Theta$.

If $A = F \cap \Theta \in \mathcal{F}_\Theta$ for some $F \in \mathcal{F}$, then $\Theta \setminus A = F^c \cap \Theta \in \mathcal{F}_\Theta$, since $F^c \in \mathcal{F}$ (note that $\Theta \setminus A$ is the complement of A in Θ).

If $\{A_i\} \subset \mathcal{F}_\Theta$, then each $A_i = F_i \cap \Theta$, for some $\{F_i\} \subset \mathcal{F}$. In this case, $\cup_{i\geq 1} F_i \in \mathcal{F}$ and hence $\cup_{i\geq 1} A_i = (\cup_{i\geq 1} F_i) \cap \Theta \in \mathcal{F}_\Theta$. This proves that $\mathcal{F}_\Theta$ is a σ-algebra of subsets of Θ.

4. (**a.**) Compute

$$\int g^r d\mu = \sum_{i\geq 1} g^r(n) = \sum_{n\geq 1} n^{\frac{-r}{p}}$$

which converges if and only if $-\frac{r}{p} < -1$ or $r > p$.
(**b.**) Similarly, compute

$$\int g^r d\mu = \sum_{n\geq 1} g^r(n) \frac{1}{n^2} = \sum_{i\geq 1} n^{\frac{r}{p}} \frac{1}{n^2}$$

which converges if and only if $\frac{r}{p} - 2 < -1$ or $r < p$.

5. (**a.**) Under the given conditions, the *strong* law of large numbers for $\{X_n\}$ says

$$\frac{S_n}{n} \xrightarrow{a.s.} 0.$$

(**b.**) This follows from the fact that a.s convergence implies convergence in probability.

(**c.**) To show this, note that if φ_n, φ are the characteristic functions of $\frac{S_n}{n}$ and X_1 respectively, then by properties of the characteristic functions, we have

$$\varphi_n(t) = E\left(\exp\left(it\frac{S_n}{n}\right)\right) = \left(\varphi\left(\frac{t}{n}\right)\right)^n, \tag{1}$$

Now note that if $\beta(t) = \varphi(t) - 1$, then

$$\frac{|\beta(t)|}{t} = \frac{|\varphi(t) - 1|}{t} \leq \min\left\{\frac{tE|X_1|^2}{2}, 2E|X|\right\} \to 0$$

as $t \to 0$ (or in other words, $\beta(t) = o(t)$ as $t \to 0$). Here we used the fact that $E(X_1) = 0$ as well as the fact that $E|X|^2 < \infty$ (and $E|X| < \infty$). This implies that (for a fixed t) for all $\epsilon > 0$, $|\beta(t/n)| \leq \epsilon/n$ for n large and hence, from (1),

$$\varphi_n(t) = (1 + \beta(t/n))^n \leq (1 + |\beta(t/n)|)^n \leq (1 + \epsilon/n)^n \to e^{\epsilon}, \text{ as } n \to \infty.$$

$$\varphi_n(t) = (1 + \beta(t/n))^n \geq (1 - |\beta(t/n)|)^n \geq (1 - \epsilon/n)^n \to e^{-\epsilon}, \text{ as } n \to \infty.$$

Since $\epsilon > 0$ is arbitrary, it implies that as $n \to \infty$, $\varphi_n(t) \to \phi_0(t) \equiv 1$ and $\phi_0(t) \equiv 1$ is the characteristic function of the 0 random variable. Hence, we get that $\frac{S_n}{n} \xrightarrow{d} 0$. Since the limit is constant, we get that this convergence holds in probability as well.

(**d.**) Note that by strong law of large numbers (under the given conditions), we have

$$\frac{T_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i^2, \quad \xrightarrow{a.s.} \quad \sigma^2 \text{ as } n \to \infty, \tag{2}$$

where $\sigma^2 = E(X_1^2) < \infty$. Similarly, using the central limit theorem, we get

$$\frac{S_n}{\sqrt{n}} \quad \xrightarrow{d} \quad \mathcal{N}(0, \sigma^2). \tag{3}$$

Hence, using Slutsky's theorem, we get the result using (2) and (3).

6. (**a.**) Define for $n \geq 1$ and a given real $b > 0$,

$$A_{n,b} = \{w \in \Omega : |b^{-1}X_n(\omega)| \geq n\} \equiv \{|X_n| \geq bn\} = \left\{\frac{|X_n|}{n} \geq b\right\}.$$

and by the two Borel-Cantelli Lemmas, we have for any independent events $\{A_{n,b}\}$

$$\sum_{n\geq 1} P(A_{n,b}) < \infty \text{ iff } P(\limsup A_{n,b}) = 0 \tag{4}$$

Now recall that

$$\sum_{n\geq 1} P(b^{-1}|X_1| \geq n) \leq E|b^{-1}X_1| \leq 1 + \sum_{n\geq 1} P(b^{-1}|X_1| \geq n) \tag{5}$$

Hence, using the fact that $\{X_n\}$ is an iid sequence, we get

$$\begin{aligned}
b^{-1}E|X_n| < \infty \quad & \text{iff} \quad \sum_{n\geq 1} P(b^{-1}|X_1| \geq n) < \infty \quad (\text{ using (5)}) \\
& \text{iff} \quad \sum_{n\geq 1} P(A_{n,b}) < \infty \quad (\text{ using } X_n\text{'s iid, } X_1, X_n \text{ has same law}) \\
& \text{iff} \quad P(\limsup A_{n,b}) = 0 \quad (\text{using (4)}) \\
& \text{iff} \quad P([\limsup A_{n,b}]^c) = 1 \quad (\text{taking the complement}) \qquad (6)
\end{aligned}$$

where the event in the last line of the display above is:

$$\begin{aligned}
& [\limsup A_{n,b}]^c \\
= \; & \left\{\omega \in \Omega: \text{ there exists } N = N(\omega), \text{ such that } \frac{|X_n(\omega)|}{n} < b, \text{ for } n \geq N\right\}
\end{aligned}$$

Hence, if $E|X_1| < \infty$, then it holds that $P([\limsup A_{n,1}]^c) = 1$ from (6) with $b = 1$; if $\omega \in \limsup A_{n,1}$, then $\sup\{|X_n(\omega)|/n : n \geq N(\omega)\} \leq 1$ for some $N(\omega)$ by (7) so that $\limsup_{n\to\infty} |X_n(\omega)|/n \leq 1 < \infty$. Hence, $P(\limsup_{n\to\infty} |X_n|/n < \infty) = 1$. For the other direction, note that $D = \limsup_{n\to\infty} |X_n|/n$ is a non-negative (possibly extended) tail random variable; by the independence of the X_n's and Kolmogorov's 0-1 Law, it follows that D must be degenerate. So if $P(D < \infty) = 1$, then there exists a real $d \geq 0$ such that $P(D = d) = 1$. By $P(D = d) = 1$, (6) and (7), it must hold that $P(\limsup A_{n,b}) = 1$ with $b = d+1$, implying $b^{-1}E|X_1| < \infty$ with $b = d + 1$.

(**b.**) Note that if $E|X_1| < \infty$ then by the strong law of large numbers, $S_n/n \to E(X_1) < \infty$ as $n \to \infty$ with probability 1. Hence, $\sup_{n\geq 1} |S_n|/n < \infty$ with probability 1 (this is ω-by-ω argument that every convergent real sequence is bounded).

For the other direction, note that

$$\frac{|X_n|}{n} \leq \frac{|S_n - S_{n-1}|}{n} \leq \frac{|S_n|}{n} + \left(\frac{n-1}{n}\right)\frac{|S_{n-1}|}{n-1} \leq 2\sup_{n\geq 1} \frac{|S_n|}{n}$$

Hence, this direction follows using part (a)