(A1, B1) (*14pts, 2pts for each*)

Exam A.
(a) False (b) True (c) False (d) True
(e) True (f) False (g) True

Exam B.
(a) False (b) False (c) True (d) False
(e) True (f) True (g) False

(A2, B2) (*12pts, 6pts for each*)

Exam A.

(a) **(1)** $(X_0, X_1, \ldots, X_5) \sim \text{multinomial}(100, q_0(p), q_1(p), \ldots, q_5(p))$, where $q_i(p)$'s are the functions of p as defined in the question, and p is the probability of getting head in tossing a coin.

(2) $\Omega = \{p : 0 < p < 1\}$.

(3) $\Omega_0 = \{p : p = 1/2\}$. If the null is rejected, it would be concluded that the coins are not fair.

(b) **(1)** The random variables X_A and X_B are independent, and $X_A \sim \text{binomial}(50, p_A)$ and $X_B \sim \text{binomial}(50, p_B)$, where p_A and p_B are the probability that a subject judges automobile 1 to be more expensive in groups A and B respectively.

(2) $\Omega = \{(p_A, p_B) : 0 < p_A < 1,\ 0 < p_B < 1\}$

(3) $\Omega_0 = \{(p_A, p_B) : 0 < p_A \le p_B < 1\}$
If the null hypothesis is rejected, it would be concluded that a female model can increase the perceived cost of an automobile.

Other acceptable answers are:

---

(2) $\Omega = \{(p_A, p_B) : 0 < p_A < 1,\ p_B = 1/2\}$

(3) $\Omega_0 = \{(p_A, p_B) : 0 < p_A \le 1/2,\ p_B = 1/2\}$

This formulation assumes that without a female model, subjects have no preference between the two automobiles (i.e., $p_B = 1/2$).

Exam B.

(b) **(1)** $(X_0, X_1, X_2, X_3) \sim \text{multinomial}(88, q_0(p), q_1(p), q_2(p), q_3(p))$, where $q_i(p)$'s are the functions of p as defined in the question, and p is the probability of getting head in tossing a coin.

(2) $\Omega = \{p : 0 < p < 1\}$.

(3) $\Omega_0 = \{p : p = 1/2\}$. If the null is rejected, it would be concluded that the coins are not fair.

(a) **(1)** The random variables X_A and X_B are independent, and $X_A \sim \text{binomial}(80, p_A)$ and $X_B \sim \text{binomial}(80, p_B)$, where p_A and p_B are the probability that a subject judges automobile 1 to be more expensive in groups A and B respectively.

(2) $\Omega = \{(p_A, p_B) : 0 < p_A < 1,\ 0 < p_B < 1\}$

(3) $\Omega_0 = \{(p_A, p_B) : 0 < p_A \le p_B < 1\}$
If the null hypothesis is rejected, it would be concluded that a female model can increase the perceived cost of an automobile.

Other acceptable answers are:

---

(2) $\Omega = \{(p_A, p_B) : 0 < p_A < 1,\ p_B = 1/2\}$

(3) $\Omega_0 = \{(p_A, p_B) : 0 < p_A \le 1/2,\ p_B = 1/2\}$

This formulation assumes that without a female model, subjects have no preference between the two automobiles (i.e., $p_B = 1/2$).

(A2, B2) (*12pts, cont.*)

Exam A.

(b) **(2)** $\Omega = \{(p_A, p_B) : 0 < p_B \leq p_A < 1\}$

(3) $\Omega_0 = \{(p_A, p_B) : 0 < p_B = p_A < 1\}$

This formulation reflects the assumption that using a female model cannot decrease the perceived cost of the automobile (i.e., $p_A \geq p_B$).

---

(2) $\Omega = \{(p_A, p_B) : 1/2 \leq p_A < 1,\ p_B = 1/2\}$

(3) $\Omega_0 = \{(p_A, p_B) : p_A = p_B = 1/2\}$

This formulation assumes that without a female model, subjects have no preference between the two automobiles (i.e., $p_B = 1/2$). It further assumes that the presence of a female model cannot decrease the perceived cost of the automobile (i.e., $p_A \geq p_B$).

Exam B.

(a) **(2)** $\Omega = \{(p_A, p_B) : 0 < p_B \leq p_A < 1\}$

(3) $\Omega_0 = \{(p_A, p_B) : 0 < p_B = p_A < 1\}$

This formulation reflects the assumption that using a female model cannot decrease the perceived cost of the automobile (i.e., $p_A \geq p_B$).

---

(2) $\Omega = \{(p_A, p_B) : 1/2 \leq p_A < 1,\ p_B = 1/2\}$

(3) $\Omega_0 = \{(p_A, p_B) : p_A = p_B = 1/2\}$

This formulation assumes that without a female model, subjects have no preference between the two automobiles (i.e., $p_B = 1/2$). It further assumes that the presence of a female model cannot decrease the perceived cost of the automobile (i.e., $p_A \geq p_B$).

(A3, B3) *(15pts)*

Exam A.

(a) *(2pts)* Both hypotheses are simple.

(b) *(4pts)* By the Neyman–Pearson lemma, the most powerful test for testing simple hypotheses rejects H_0 when the likelihood ratio

$$\Lambda(x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n \mid \theta = 1)}{f(x_1, \ldots, x_n \mid \theta = 2)}$$

is sufficiently small.

The joint density of $X_1, \ldots, X_n \sim$ i.i.d. uniform$(0, \theta)$ is

$$f(x_1, \ldots, x_n \mid \theta) = \frac{1}{\theta^n} \cdot \mathbf{1}_{\{X_{(n)} \le \theta\}}.$$

Thus, the likelihood ratio becomes

$$\Lambda(x_1, \ldots, x_n) = \begin{cases} 2^n, & \text{if } 0 \le x_{(n)} \le 1 \\ 0, & \text{if } 1 < x_{(n)} \le 2 \end{cases}$$

Hence, the most powerful test at level $\alpha = 0$ is:

$$\text{reject } H_0 \text{ if } X_{(n)} > 1$$

because under H_0, $\alpha = P(X_{(n)} > 1 \mid \theta = 1) = 0$.

Under $H_A : \theta = 2$,

$$P(X_{(n)} > 1 \mid \theta = 2) = 1 - \left(\frac{1}{2}\right)^n = 1 - 2^{-n}$$

Therefore, the power is $1 - 2^{-n}$.

(c) *(2pts)* Under $H_0 : \theta = 1$, the significance level is

$$\alpha = P(a \le X_{(n)} \le 1] \mid \theta = 1) = 1 - a^n.$$

Under $H_A : \theta = 2$, the power is

$$P(a \le X_{(n)} \le 2 \mid \theta = 2) = 1 - \left(\frac{a}{2}\right)^n.$$

Exam B.

(a) *(2pts)* Both hypotheses are simple.

(b) *(4pts)* By the Neyman–Pearson lemma, the most powerful test for testing simple hypotheses rejects H_0 when the likelihood ratio

$$\Lambda(x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n \mid \theta = 1)}{f(x_1, \ldots, x_n \mid \theta = 3)}$$

is sufficiently small.

The joint density of $X_1, \ldots, X_n \sim$ i.i.d. uniform$(0, \theta)$ is

$$f(x_1, \ldots, x_n \mid \theta) = \frac{1}{\theta^n} \cdot \mathbf{1}_{\{X_{(n)} \le \theta\}}.$$

Thus, the likelihood ratio becomes

$$\Lambda(x_1, \ldots, x_n) = \begin{cases} 3^n, & \text{if } 0 \le x_{(n)} \le 1 \\ 0, & \text{if } 1 < x_{(n)} \le 3 \end{cases}$$

Hence, the most powerful test at level $\alpha = 0$ is:

$$\text{reject } H_0 \text{ if } X_{(n)} > 1$$

because under H_0, $\alpha = P(X_{(n)} > 1 \mid \theta = 1) = 0$.

Under $H_A : \theta = 3$,

$$P(X_{(n)} > 1 \mid \theta = 3) = 1 - \left(\frac{1}{3}\right)^n = 1 - 3^{-n}$$

Therefore, the power is $1 - 3^{-n}$.

(c) *(2pts)* Under $H_0 : \theta = 1$, the significance level is

$$\alpha = P(a \le X_{(n)} \le 1] \mid \theta = 1) = 1 - a^n.$$

Under $H_A : \theta = 3$, the power is

$$P(a \le X_{(n)} \le 3 \mid \theta = 3) = 1 - \left(\frac{a}{3}\right)^n.$$

(A3, B3) (*15pts*, cont.)

Exam A.

(d) (*3pts*) Under $H_0 : \theta = 1$, $P(X_{(n)} > 1 \mid \theta = 1) = 0$ and $P(X_{(n)} \le 1 \mid \theta = 1) = 1$, so the significance level of the randomized test is

$$\alpha = \gamma \times P(X_{(n)} \le 1 \mid \theta = 1) = \gamma \times 1 = \gamma.$$

Under $H_A : \theta = 2$, we compute:

$$P(X_{(n)} > 1 \mid \theta = 2) = 1 - \left(\frac{1}{2}\right)^n = 1 - 2^{-n},$$

and $P(X_{(n)} \le 1 \mid \theta = 2) = 2^{-n}$. Therefore, the power is

$$1 - 2^{-n} + \gamma \cdot 2^{-n}.$$

(e) (*4pts*) The most powerful test of level $\alpha \in (0,1)$ is not unique. Consider the following example. Let $a = a(\alpha) = (1-\alpha)^{1/n}$. Then the non-randomized test from part (c) that rejects H_0 when $X_{(n)} \in [a, 2]$ has significance level α. Its power is

$$1 - \left(\frac{a}{2}\right)^n = 1 - \frac{1-\alpha}{2^n} = 1 - 2^{-n}(1-\alpha).$$

Now consider the randomized test from part (d) with $\gamma = \alpha$. Its significance level is also α, and its power is

$$1 - 2^{-n}(1-\gamma) = 1 - 2^{-n}(1-\alpha).$$

Therefore, both tests have the same level α, and the same power, but different rejection rules (non-randomized vs. randomized). Thus, the most powerful test is not unique.

Exam B.

(d) (*3pts*) Under $H_0 : \theta = 1$, $P(X_{(n)} > 1 \mid \theta = 1) = 0$ and $P(X_{(n)} \le 1 \mid \theta = 1) = 1$, so the significance level of the randomized test is

$$\alpha = \gamma \times P(X_{(n)} \le 1 \mid \theta = 1) = \gamma \times 1 = \gamma.$$

Under $H_A : \theta = 3$, we compute:

$$P(X_{(n)} > 1 \mid \theta = 3) = 1 - \left(\frac{1}{3}\right)^n = 1 - 3^{-n},$$

and $P(X_{(n)} \le 1 \mid \theta = 3) = 3^{-n}$. Therefore, the power is

$$1 - 3^{-n} + \gamma \cdot 3^{-n}.$$

(e) (*4pts*) The most powerful test of level $\alpha \in (0,1)$ is not unique. Consider the following example. Let $a = a(\alpha) = (1-\alpha)^{1/n}$. Then the non-randomized test from part (c) that rejects H_0 when $X_{(n)} \in [a, 3]$ has significance level α. Its power is

$$1 - \left(\frac{a}{3}\right)^n = 1 - \frac{1-\alpha}{3^n} = 1 - 3^{-n}(1-\alpha).$$

Now consider the randomized test from part (d) with $\gamma = \alpha$. Its significance level is also α, and its power is

$$1 - 3^{-n}(1-\gamma) = 1 - 3^{-n}(1-\alpha).$$

Therefore, both tests have the same level α, and the same power, but different rejection rules (non-randomized vs. randomized). Thus, the most powerful test is not unique.

(A4, B4) *(14pts)*

Exam A.

(a) *(10pts)* The joint pmf of $X_1, \ldots, X_n$ is

$$f(\mathbf{x}, \lambda) = e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i} \cdot \left(\prod_{i=1}^n \frac{1}{x_i!} \right).$$

Therefore, the likelihood ratio is

$$\Lambda(\mathbf{x}) = \frac{f(\mathbf{x}, \lambda_0)}{f(\mathbf{x}, \lambda_1)} = e^{n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_0}{\lambda_1} \right)^{\sum_{i=1}^n x_i}.$$

Because $\lambda_1 > \lambda_0 \Leftrightarrow \lambda_0/\lambda < 1$, Λ decreases as $\sum_{i=1}^n X_i$ increases. A randomized test function based on the likelihood ratio is then given by:

$$\phi(\mathbf{x}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i > c \\ \gamma, & \text{if } \sum_{i=1}^n X_i = c \\ 0, & \text{if } \sum_{i=1}^n X_i < c \end{cases}.$$

Under $H_0 : \lambda = \lambda_0$, the c and γ are determined by

$$\begin{aligned} &\mathrm{E}_{\lambda_0}(\phi) \\ &= P\left(\sum_{i=1}^n X_i > c \right) + \gamma \cdot P\left(\sum_{i=1}^n X_i = c \right) \\ &= \alpha, \end{aligned}$$

where $\sum_{i=1}^n X_i \sim P(n\lambda_0)$.

(b) *(4pts)* By Neyman-Pearson lemma, the test in (a) is the most powerful test for any particular simple alternative $H_A : \lambda = \lambda_1$, where $\lambda_1 > \lambda_0$. Furthermore, because the rejection region, i.e., c and γ, of the test does not depend on λ_1, the test is UMP for $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda > \lambda_0$.

Exam B.

(a) *(10pts)* The joint pmf of $X_1, \ldots, X_n$ is

$$f(\mathbf{x}, \lambda) = e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i} \cdot \left(\prod_{i=1}^n \frac{1}{x_i!} \right).$$

Therefore, the likelihood ratio is

$$\Lambda(\mathbf{x}) = \frac{f(\mathbf{x}, \lambda_0)}{f(\mathbf{x}, \lambda_1)} = e^{n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_0}{\lambda_1} \right)^{\sum_{i=1}^n x_i}.$$

Because $\lambda_1 < \lambda_0 \Leftrightarrow \lambda_0/\lambda > 1$, Λ decreases as $\sum_{i=1}^n X_i$ decreases. A randomized test function based on the likelihood ratio is then given by:

$$\phi(\mathbf{x}) = \begin{cases} 1, & \text{if } \sum_{i=1}^n X_i < c \\ \gamma, & \text{if } \sum_{i=1}^n X_i = c \\ 0, & \text{if } \sum_{i=1}^n X_i > c \end{cases}.$$

Under $H_0 : \lambda = \lambda_0$, the c and γ are determined by

$$\begin{aligned} &\mathrm{E}_{\lambda_0}(\phi) \\ &= P\left(\sum_{i=1}^n X_i < c \right) + \gamma \cdot P\left(\sum_{i=1}^n X_i = c \right) \\ &= \alpha, \end{aligned}$$

where $\sum_{i=1}^n X_i \sim P(n\lambda_0)$.

(b) *(4pts)* By Neyman-Pearson lemma, the test in (a) is the most powerful test for any particular simple alternative $H_A : \lambda = \lambda_1$, where $\lambda_1 < \lambda_0$. Furthermore, because the rejection region, i.e., c and γ, of the test does not depend on λ_1, the test is UMP for $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda < \lambda_0$.

(A5, B6) *(28pts)*

(a) *(7pts)* Because $\Omega_0 = \{\theta_0\}$ and $\Omega = (0, \infty)$,

$$\Lambda(x_1, \dots, x_n) = \frac{\max_{\theta \in \Omega_0} \mathcal{L}(\theta, \mathbf{x})}{\max_{\theta \in \Omega} \mathcal{L}(\theta, \mathbf{x})} = \frac{\left[\dfrac{1}{\theta_0}\right]^n \cdot I_{[0,\theta_0]}(x_{(n)})}{\left[\dfrac{1}{x_{(n)}}\right]^n \cdot I_{[0,x_{(n)}]}(x_{(n)})} = \left[\frac{x_{(n)}}{\theta_0}\right]^n \cdot I_{[0,\theta_0]}(x_{(n)}).$$

(b) *(4pts)* As a function of $x_{(n)}$, Λ increases from 0 to 1 when $x_{(n)}$ increases from 0 to θ_0, and $\Lambda = 0$ when $x_{(n)}$ is larger than θ_0. Therefore, for $0 < s < 1$,

$$\begin{aligned} \Lambda(X_1, \dots, X_n) > s \quad &\Leftrightarrow \quad X_{(n)} < \theta_0 \text{ and } \left[\frac{X_{(n)}}{\theta_0}\right]^n > s \\ &\Leftrightarrow \quad s^{1/n}\theta_0 < X_{(n)} < \theta_0. \end{aligned}$$

(c) *(8pts)* The rejection region of the GLR test is $\Lambda < s$, i.e.,

$$\{X_{(n)} < s^{1/n}\theta_0\} \cup \{X_{(n)} > \theta_0\}, \tag{I}$$

where s is determined by

$$\begin{aligned} \alpha &= P(\Lambda < s | H_0) = P\left(\{X_{(n)} > \theta_0\} \cup \{X_{(n)} < s^{1/n}\theta_0\} \middle| H_0\right) \\ &= P\left(X_{(n)} > \theta_0 \middle| \theta = \theta_0\right) + P\left(X_{(n)} < s^{1/n}\theta_0 \middle| \theta = \theta_0\right) \\ &= \int_{\theta_0}^{\infty} 0 \, dx + \int_0^{s^{1/n}\theta_0} \frac{nx^{n-1}}{\theta_0^n} \, dx = 0 + \left.\frac{x^n}{\theta_0^n}\right|_0^{s^{1/n}\theta_0} = s. \end{aligned}$$

Therefore, for $\alpha = 0.05$, we can substitute $s = 0.05$ into (I) to get the rejection region $\left\{X_{(n)} < (0.05)^{1/n}\theta_0\right\} \cup \left\{X_{(n)} > \theta_0\right\}$.

(d) *(4pts)* The acceptance region is $(0.05)^{1/n}\theta_0 \le X_{(n)} \le \theta_0$, i.e.,

$$\begin{aligned} 0.95 &= P\left((0.05)^{1/n}\theta_0 \le X_{(n)} \le \theta_0 \middle| H_0\right) \\ &= P\left(\frac{(0.05)^{1/n}}{X_{(n)}} \le \frac{1}{\theta_0} \le \frac{1}{X_{(n)}} \middle| \theta = \theta_0\right) \\ &= P\left(X_{(n)} \le \theta_0 \le \frac{X_{(n)}}{(0.05)^{1/n}} \middle| \theta = \theta_0\right) \end{aligned}$$

Therefore, $\left[X_{(n)}, 20^{1/n}X_{(n)}\right]$ is a 95% confidence interval for θ.

(e) *(5pts)* Using the equality from (d):

$$0.05 = P\left((0.05)^{1/n} \le \frac{X_{(n)}}{\theta} \le 1 \middle| \theta\right) \quad \text{for any } \theta,$$

we can identify a pivotal quantity $X_{(n)}/\theta$, which is a function of data and parameter. It is really a pivotal quantity because its pdf and/or cdf, which are respectively

$$\begin{aligned} f_{X_{(n)}/\theta}(x) &= nx^{n-1}, \quad \text{for } 0 \le x \le 1, \\ F_{X_{(n)}/\theta}(x) &= x^n, \quad \text{for } 0 \le x \le 1, \end{aligned}$$

are irrelevant to θ. Alternatively, $X_{(n)}/\theta = \max\{X_1/\theta, \dots, X_n/\theta\}$, where $X_1/\theta, \dots, X_n/\theta \sim$ i.i.d. uniform(0, 1) have a joint distribution irrelevent to θ.

(A6, B5) *(17pts)*

(a) *(5pts)* The posterior pdf is

$$\begin{aligned}
h(\theta|x_1,\dots,x_n) &\propto f(x_1,\dots,x_n|\theta)\cdot g(\theta) = \left[\prod_{i=1}^{n} f(x_i|\theta)\right]\cdot g(\theta) \\
&\propto \left[\prod_{i=1}^{n} \theta e^{-\theta x_i}\right]\cdot \theta^{\alpha-1}e^{-\lambda\theta} \\
&= \theta^n e^{-\theta(\sum_{i=1}^{n} x_i)}\cdot \theta^{\alpha-1}e^{-\lambda\theta} \\
&= \theta^{(\alpha+n)-1}e^{-(\lambda+n\overline{X})\theta},
\end{aligned}$$

which follows the form of the pdf of gamma distribution with shape parameter $\alpha+n$ and scale parameter $\lambda+n\overline{X}$, i.e., $\Theta|x_1,\dots,x_n \sim \Gamma(\alpha+n,\lambda+n\overline{X})$.

(b) *(5pts)* Because $\Theta|x_1,\dots,x_n \sim \text{gamma}(\alpha+n,\lambda+n\overline{X})$, under squared error loss the Bayes estimator is

$$\hat{\theta}_B = E\left(\Theta|x_1,\dots,x_n\right) = \frac{\alpha+n}{\lambda+n\overline{X}} = \frac{\lambda}{\lambda+n\overline{X}}\cdot\frac{\alpha}{\lambda} + \frac{n\overline{X}}{\lambda+n\overline{X}}\cdot\frac{1}{\overline{X}}, \tag{II}$$

where $\frac{\alpha}{\lambda}$ is the prior mean, $\frac{1}{\overline{X}}$ is $\hat{\theta}_{\text{MLE}}$, and sum of the weights is one, i.e., $\frac{\lambda}{\lambda+n\overline{X}}+\frac{n\overline{X}}{\lambda+n\overline{X}} = 1$.

(c) *(3pts)* Because $\overline{X}$ will converge in probability to a constant according to the law of large number, when n is large, the weights in (II) will approximate 0 and 1 respectively, i.e.,

$$\frac{\lambda}{\lambda+n\overline{X}} \approx 0 \quad \text{and} \quad \frac{n\overline{X}}{\lambda+n\overline{X}} \approx 1.$$

Therefore, $\hat{\theta}_B \approx 1/\overline{X} = \hat{\theta}_{\text{MLE}}$, which is a function of sample (data) only and is irrelevent to the prior.

(d) *(4pts)* For some values of θ, especially those far from the prior mean $\frac{\alpha}{\lambda}$, the Bayes estimator $\hat{\theta}_B$ may have enough bias such that its mean squared error exceeds that of the MLE $\hat{\theta}_{\text{MLE}}$. However, the $\hat{\theta}_{\text{MLE}}$ *cannot strictly dominate* the $\hat{\theta}_B$ across all $\theta > 0$, because Bayes estimators are *admissible*, i.e., there does not exist any estimator, including the $\hat{\theta}_{\text{MLE}}$, whose risk function is no larger than that of $\hat{\theta}_B$ at every value of θ, and strictly smaller at some value of θ. Therefore, strict domination by the $\hat{\theta}_{\text{MLE}}$ is impossible.