1. (*14pts, 2pts for each*)
(a) True (b) False (c) False (d) True (e) False ~~(f) False (g) True~~

2. (*12pts, 4pts for each*)

(a) Let N be the total number of objects that had been manufactured by the company. Then, $X_1, \ldots, X_n$ are i.i.d. $\sim$ uniform$(1, 2, \ldots, N)$ (they are independent because of sampling with replacement), where N is an unknown parameter.

(b) Let p be the probability of obtaining a head when the coin is tossed, then $X_1 \sim$ binomial$(3, p)$ and $X_2 \sim$ geometric(p) (or negative binomial$(1, p)$), and X_1 and X_2 are independent. The probability p is an unknown parameter.

(c) Let p_W be the proportion of patients in the population whose race is White and other, and let p_1 be the proportion of patients in the population who had been advised by the physicians to stop smoking. Because whether a patient was advised by the physicians is independent of the race of the patient, the probabilities that a patient falls in each of the 4 categories $(W, 1)$, $(W, 2)$, $(A, 1)$, and $(A, 2)$ are $p_W p_1$, $p_W(1 - p_1)$, $(1 - p_W)p_1$, and $(1 - p_W)(1 - p_1)$, respectively. Then, $(X_{W,1}, X_{W,2}, X_{A,1}, X_{A,2}) \sim$ multinomial$(311, 4, p_W p_1, p_W(1 - p_1), (1 - p_W)p_1, (1 - p_W)(1 - p_1))$, where the probabilities p_W and p_1 are unknown parameters in the model.

3. (*18pts*)

(a) (*3pts*) Because for $i = 1, \ldots, n$,

$$\mu = \mathrm{E}(Y_i) = \theta + \frac{1}{2} \Rightarrow \theta = \mu - \frac{1}{2},$$

the moment estimator is $\hat{\theta}_1 = \overline{Y} - \frac{1}{2}$. The estimator $\hat{\theta}_1$ is unbiased because

$$\mathrm{E}(\hat{\theta}_1) = \mathrm{E}(\overline{Y} - \frac{1}{2}) = \mathrm{E}(\overline{Y}) - \frac{1}{2} = \frac{n\mu}{n} - \frac{1}{2} = \left(\theta + \frac{1}{2}\right) - \frac{1}{2} = \theta.$$

(b) (*2pts*) The variance of $\hat{\theta}_1$ is

$$\mathrm{Var}(\hat{\theta}_1) = \mathrm{Var}\left(\overline{Y} - \frac{1}{2}\right) = \mathrm{Var}(\overline{Y}) = \frac{1}{n}\mathrm{Var}(Y_1) = \frac{1}{n} \times \frac{[(\theta + 1) - \theta]^2}{12} = \frac{1}{12n}.$$

The standard error of $\hat{\theta}_1$ is $\frac{1}{\sqrt{12n}}$.

(c) (*4pts*) From the hints (i)(ii) and Thm 2.8 in LN, Ch1-6, p.35, the pdf of $T_{(n)}$ is

$$f(t) = nt^{n-1}, \quad 0 < t < 1. \tag{I}$$

The mean of $T_{(n)}$ is

$$\mathrm{E}(T_{(n)}) = \int_0^1 t \cdot nt^{n-1} dt = \frac{n}{n+1}, \tag{II}$$

and $\hat{\theta}_2$ is unbiased because

$$\mathrm{E}(\hat{\theta}_2) = \mathrm{E}(Y_{(n)} - \theta) + \theta - \frac{n}{n+1} = \mathrm{E}(T_{(n)}) + \theta - \frac{n}{n+1} = \theta.$$

(d) (*4pts*) From equation (I),

$$\mathrm{E}\left(T_{(n)}^2\right) = \int_0^1 t^2 \cdot nt^{n-1} dt = \frac{n}{n+2}. \tag{III}$$

From hint (ii) and equations (II) and (III),

$$\mathrm{Var}(\hat{\theta}_2) = \mathrm{Var}(Y_{(n)}) = \mathrm{Var}(T_{(n)}) = \mathrm{E}(T_{(n)}^2) - (\mathrm{E}(T_{(n)}))^2 = \frac{n}{(n+1)^2(n+2)}.$$

The standard error of $\hat{\theta}_2$ is $\sqrt{\frac{n}{(n+1)^2(n+2)}}$.

(e) (*2pts*) The relative efficiency is:

$$\mathrm{eff}_n(\hat{\theta}_1, \hat{\theta}_2) = \frac{\mathrm{Var}(\hat{\theta}_2)}{\mathrm{Var}(\hat{\theta}_1)} = \frac{12n^2}{(n+1)^2(n+2)}.$$

The asymptotic relative efficiency is

$$\lim_{n\to\infty} \mathrm{eff}_n(\hat{\theta}_1, \hat{\theta}_2) = 0,$$

which shows $\hat{\theta}_2$ has a much smaller variance than $\hat{\theta}_1$ when the sample size is large.

(f) (*3pts*) Because $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased, $MSE(\hat{\theta}_i) = \mathrm{Var}(\hat{\theta}_i)$, for $i = 1, 2$. We know that $\mathrm{eff}_n(\hat{\theta}_1, \hat{\theta}_2) < 1$ for $n > 7$, which implies $\mathrm{Var}(\hat{\theta}_1) > \mathrm{Var}(\hat{\theta}_2)$ when the sample size is greater than 7. In the case, $\hat{\theta}_2$ has a smaller MSE and is better.

4. (*17pts*)

(a) (*4pts*) The joint pdf is:

$$\prod_{i=1}^n f(y_i|\theta) = \frac{1}{\theta^n} r^n \left(\prod_{i=1}^n y_i\right)^{r-1} e^{-\frac{\sum_{i=1}^n y_i^r}{\theta}} = g(t,\theta)h(y_1,\dots,y_n),$$

where

$$t = \sum_{i=1}^n y_i^r, \quad g(t,\theta) = \frac{1}{\theta^n} e^{-\frac{t}{\theta}}, \quad \text{and} \quad h(y_1,\dots,y_n) = r^n \left(\prod_{i=1}^n y_i\right)^{r-1}.$$

By the factorization theorem, the statistic $T = \sum_{i=1}^n Y_i^r$ is sufficient.

(b) (*4pts*) The log-likelihood function is:

$$l(\theta) = -n\log(\theta) - \frac{\sum_{i=1}^n y_i^r}{\theta} + \log(h(y_1,\dots,y_n)).$$

By solving the normal equation:

$$l'(\theta) = -\frac{n}{\theta} + \frac{\sum_{i=1}^n y_i^r}{\theta^2} = 0,$$

we obtain the solution: $\frac{\sum_{i=1}^n y_i^r}{n}$. Because

$$l''(\theta) = -\frac{n}{\theta^2} - \frac{2\sum_{i=1}^n y_i^r}{\theta^3} < 0,$$

for any θ, this solution is MLE.

(c) (*4pts*) Let $X = Y^r$, then $Y = X^{\frac{1}{r}}$ and $X > 0$. Because $|J| = |\frac{dy}{dx}| = \frac{1}{r}x^{\frac{1}{r}-1}$, for $x > 0$, the pdf of X is

$$f(x|\theta) = \left(\frac{1}{\theta}\right) r x^{\frac{r-1}{r}} e^{-\frac{x}{\theta}} \times \frac{1}{r} x^{\frac{1}{r}-1} = \frac{1}{\theta} e^{-\frac{x}{\theta}},$$

which is the pdf of exponential$(1/\theta)$.

(d) (*5pts*) Because

$$\frac{d^2}{d\theta^2}\log(f(y|\theta)) = \frac{1}{\theta^2} - \frac{2}{\theta^3}y^r,$$

and $\mathrm{E}(Y^r) = \mathrm{E}(X) = \theta$ (from (c)), we can get the Fisher information of one observation:

$$I(\theta) = -\mathrm{E}\left(\frac{1}{\theta^2} - \frac{2}{\theta^3}Y^r\right) = -\frac{1}{\theta^2} + \frac{2}{\theta^3}\mathrm{E}(Y^r) = -\frac{1}{\theta^2} + \frac{2}{\theta^3}\theta = \frac{1}{\theta^2}.$$

The Fisher information contained in $Y_1, \ldots, Y_n$ is $nI(\theta) = n/\theta^2$, and the asymptotic variance of the MLE is

$$\frac{1}{nI(\theta)} = \frac{\theta^2}{n}.$$

5. (*27pts*)

(a) (*4pts*) Because

$$P(z|\text{identical twins}) = \begin{cases} 1/2, & \text{if } z = MM, \\ 1/2, & \text{if } z = FF, \\ 0, & \text{if } z = MF, \end{cases}$$

and

$$P(z|\text{non-identical twins}) = \begin{cases} 1/4, & \text{if } z = MM, \\ 1/4, & \text{if } z = FF, \\ 1/2, & \text{if } z = MF, \end{cases}$$

by the law of total probability, we have

$$\begin{aligned} P(MM) &= P(\text{identical twins})P(MM|\text{identical twins}) \\ &\quad + P(\text{non-identical twins})P(MM|\text{non-identical twins}) \\ &= \alpha \times (1/2) + (1-\alpha) \times (1/4) = (1+\alpha)/4. \end{aligned}$$

Similarly, we can get $P(FF) = (1+\alpha)/4$ and $P(MF) = (1-\alpha)/2$.

(b) (*2pts*) $(X_1, X_2, X_3) \sim \text{multinomial}\big(n, 3, \frac{1+\alpha}{4}, \frac{1+\alpha}{4}, \frac{1-\alpha}{2}\big)$.

(c) (*4pts*) The joint pmf of (X_1, X_2, X_3) is

$$\begin{aligned} &\binom{n}{x_1\ x_2\ x_3}\left(\frac{1+\alpha}{4}\right)^{x_1+x_2}\left(\frac{1-\alpha}{2}\right)^{x_3} \\ &= \exp\left[(x_1+x_2)\log\left(\frac{1+\alpha}{4}\right) + (n - x_1 - x_2)\log\left(\frac{1-\alpha}{2}\right)\right]\binom{n}{x_1\ x_2\ x_3} \\ &= \exp\left[(x_1+x_2)\log\left(\frac{1+\alpha}{2(1-\alpha)}\right) + n\log\left(\frac{1-\alpha}{2}\right)\right]\binom{n}{x_1\ x_2\ x_3}, \end{aligned} \tag{IV}$$

for $x_i = 0, \ldots, n$, $i = 1, 2, 3$, and $x_1 + x_2 + x_3 = n$. This is a one-parameter exponential family with $c(\alpha) = \log\left(\frac{1+\alpha}{2(1-\alpha)}\right)$ and $T(x_1, x_2, x_3) = x_1 + x_2$. So, $X_1 + X_2$ is a sufficient and complete statistic for α.

(d) (*4pts*) From equation (IV), we can get the log-likelihood of (X_1, X_2, X_3):

$$l(\alpha) = (x_1 + x_2)\log\left(\frac{1+\alpha}{4}\right) + x_3 \log\left(\frac{1-\alpha}{2}\right) + \log\left(\binom{n}{x_1\ x_2\ x_3}\right).$$

The normal equation is

$$\begin{aligned} 0 = l'(\alpha) &= \frac{x_1+x_2}{1+\alpha} - \frac{x_3}{1-\alpha} = \frac{(x_1+x_2)(1-\alpha) - x_3(1+\alpha)}{1-\alpha^2} \\ &= \frac{(x_1+x_2-x_3) - (x_1+x_2+x_3)\alpha}{1-\alpha^2} = \frac{(x_1+x_2-x_3) - n\alpha}{1-\alpha^2}, \end{aligned} \tag{V}$$

which gives the solution (i.e., MLE)

$$\hat{\alpha} = \frac{X_1 + X_2 - X_3}{n}.$$

It can be easily checked that $l''(\alpha) < 0$ at $\alpha = \hat{\alpha}$ (by the equation (VI) in the solution of (g).)

(e) (*2pts*) The MLE $\hat{\alpha}$ is unbiased because

$$\begin{aligned} E(\hat{\alpha}) &= \frac{E(X_1) + E(X_2) - E(X_3)}{n} \\ &= \frac{n(1+\alpha)/4 + n(1+\alpha)/4 - n(1-\alpha)/2}{n} = \alpha. \end{aligned}$$

(f) (*4pts*) The variance of $\hat{\alpha}$ is

$$\begin{aligned} \mathrm{Var}(\hat{\alpha}) &= \frac{\mathrm{Var}(X_1+X_2-X_3)}{n^2} = \frac{\mathrm{Var}(n-2X_3)}{n^2} = \frac{4\mathrm{Var}(X_3)}{n^2} \\ &= \frac{4}{n^2} \times n\left(\frac{1-\alpha}{2}\right)\left(1 - \frac{1-\alpha}{2}\right) = \frac{1-\alpha^2}{n}. \end{aligned}$$

(g) (*5pts*) From equation (V), we can get the second derivative of log-likelihood $l(\alpha)$:

$$l''(\alpha) = \frac{-n(1-\alpha^2) + (n\hat{\alpha} - n\alpha)(-2\alpha)}{(1-\alpha^2)^2} = -\frac{n}{1-\alpha^2} - \frac{2n\alpha(\hat{\alpha}-\alpha)}{(1-\alpha^2)^2}. \tag{VI}$$

The Fisher information contained in (X_1, X_2, X_3) is

$$\begin{aligned} -\mathrm{E}[l''(\alpha)] &= \frac{n}{1-\alpha^2} + \frac{2n\alpha\left[\mathrm{E}(\hat{\alpha}) - \alpha\right]}{(1-\alpha^2)^2} \\ &= \frac{n}{1-\alpha^2} + \frac{2n\alpha(\alpha-\alpha)}{(1-\alpha^2)^2} \quad \text{(because } \hat{\alpha} \text{ is unbiased)} \\ &= \frac{n}{1-\alpha^2}. \end{aligned}$$

So, the Cramer-Rao lower bound is $\frac{1}{-\mathrm{E}[l''(\alpha)]} = \frac{1-\alpha^2}{n}$.

(h) (*2pts*) No unbiased estimators can have a variance smaller than the Cramer-Rao lower bound. Because $\hat{\alpha}$ is an unbiased estimator and its variance achieves the Cramer-Rao lower bound (i.e., $(1-\alpha^2)/n$), $\hat{\alpha}$ has the smallest variance among all unbiased estimators.

6. *(12pts)*

(a) *(2pts)* For $\theta < y < \infty$,

$$
\begin{aligned}
F_Y(y|\theta) &= P(Y \le y) \\
&= 1 - P(X_1 > y, \ldots, X_n > y) \\
&= 1 - \prod_{i=1}^{n} P(X_i > y) \\
&= 1 - \left[e^{-(y-\theta)}\right]^n = 1 - e^{-n(y-\theta)}.
\end{aligned}
$$

(b) *(4pts)* For $0 < t < \infty$, the cdf of $T_n(\theta)$ is:

$$
\begin{aligned}
P\left(T_n(\theta) \le t\right) &= P\left(2n(Y-\theta) \le t\right) \\
&= P\left(Y \le \frac{t}{2n} + \theta\right) \\
&= 1 - e^{-n\left(\frac{t}{2n}+\theta-\theta\right)} = 1 - e^{-\frac{t}{2}},
\end{aligned}
$$

which is irrelevant to θ. Because $T_n(\theta)$ is a function of parameter θ and data $X_1, \ldots, X_n$, and the distribution of $T_n(\theta)$ is irrelevant to θ, $T_n(\theta)$ is a pivotal quantity. [**Note**: The distribution of $T_n(\theta)$ is exponential$\left(\frac{1}{2}\right)$.]

(c) *(6pts)* Let a and b satisfy:

$$
\begin{aligned}
\frac{\alpha}{2} &= P\left(E\left(\frac{1}{2}\right) < a\right) = 1 - e^{-\frac{a}{2}} \\
&\Rightarrow a = -2\log\left(1 - \frac{\alpha}{2}\right), \\
\frac{\alpha}{2} &= P\left(E\left(\frac{1}{2}\right) > b\right) = e^{-\frac{b}{2}} \\
&\Rightarrow b = -2\log\left(\frac{\alpha}{2}\right).
\end{aligned}
$$

Then, from (b), we know

$$
\begin{aligned}
1-\alpha &= P\left(a \le 2n(Y-\theta) \le b\right) \\
&= P\left(Y - \frac{b}{2n} \le \theta \le Y - \frac{a}{2n}\right).
\end{aligned}
$$

Therefore,

$$
\left[Y + \frac{\log(\alpha/2)}{n}, Y + \frac{\log(1-\alpha/2)}{n}\right]
$$

is a $100(1-\alpha)\%$ confidence interval for θ.