#4 (a) We set

$$\overline{X} = E(X)$$

& solve for θ

$$E(X) = \int_{-1}^{1} x \ \frac{1}{2}(1+\theta x) \, dx$$

$$= \frac{1}{2}\int_{-1}^{1} (x + \theta x^2) \, dx$$

$$= \frac{1}{2}\left[\frac{1}{2}x^2 + \frac{\theta}{3}x^3 \Big|_{-1}^{1} \right]$$

$$= \frac{1}{2}\left[\frac{1}{2} + \frac{\theta}{3} - \frac{1}{2} + \frac{\theta}{3} \right] = \frac{\theta}{3}$$

$$\rightarrow \quad \overline{X} \stackrel{set}{=} \frac{\theta}{3} \implies \tilde{\theta} = 3\overline{X}.$$

Now,

$$E_\theta(\tilde{\theta}) = E_\theta(3\overline{X})$$

$$= 3E_\theta(\overline{X})$$

$$= 3E_\theta(X) = 3\left(\frac{\theta}{3}\right) = \theta.$$

$$\mathrm{var}_\theta(\tilde{\theta}) = \mathrm{var}_\theta(3\overline{X})$$

$$= 9\,\mathrm{var}_\theta(\overline{X}) = 9\,\frac{\mathrm{var}_\theta(X)}{n}.$$

Calculate $\mathrm{var}_\theta(X) = E_\theta(X^2) - \left(\frac{\theta}{3}\right)^2$

$$E_\theta(X^2) = \int_{-1}^{1} x^2 \ \frac{1}{2}(1+\theta x) \, dx$$

$$= \frac{1}{2}\int_{-1}^{1} (x^2 + \theta x^3)\, dx$$

$$= \frac{1}{2}\left[\frac{1}{3}x^3 + \frac{\theta}{4}x^4 \Big|_{-1}^{1} \right]$$

$$= \frac{1}{2}\left[\frac{1}{3} + \frac{\theta}{4} + \frac{1}{3} - \frac{\theta}{4} \right] = \frac{1}{3}$$

$$\Rightarrow \operatorname{var}_\theta(X) = \frac{1}{3} - \left(\frac{\theta}{3}\right)^2 = \frac{1}{3} - \frac{\theta^2}{9}$$

$$= \frac{3-\theta^2}{9}$$

$$\Rightarrow \operatorname{var}_\theta(\hat{\theta}) = 9\left(\frac{3-\theta^2}{9}\right)\Big/ n$$

$$= \frac{3-\theta^2}{n}.$$

(b) The likelihood function is

$$L(\theta \mid \underset{\sim}{x}) = \prod_{i=1}^{n} \frac{1}{2}(1+\theta x_i)\, I(-1 < x_i < 1)$$

$$= \left(\frac{1}{2}\right)^n \prod_{i=1}^{n} (1+\theta x_i)\, I(-1 < x_i < 1).$$

The log-likelihood function (for $-1 < x_i < 1$ $\forall i$) is

$$\log L(\theta \mid \underset{\sim}{x}) = -n \log 2 + \sum_{i=1}^{n} \log(1+\theta x_i)$$

The score function is

$$S(\theta \mid \underset{\sim}{x}) = \frac{\partial}{\partial\theta} \log L(\theta \mid \underset{\sim}{x})$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log(1+\theta x_i)$$

$$= \sum_{i=1}^{n} \frac{x_i}{1+\theta x_i}$$

Score equation:

$$\sum_{i=1}^{n} \frac{x_i}{1+\theta x_i} = 0 .$$

(c) We have

$$f_X(x \mid \theta) = \frac{1}{2}(1+\theta x)$$

$$\Rightarrow \log f_X(x \mid \theta) = -\ln 2 + \log(1+\theta x)$$

$$\Rightarrow \frac{\partial \log f_X(x \mid \theta)}{\partial\theta} = \frac{x}{1+\theta x}$$

$$\Rightarrow \frac{\partial^2 \log f_X(x \mid \theta)}{\partial\theta^2} = \frac{0(1+\theta x) - x(x)}{(1+\theta x)^2}$$

$$= \frac{-x^2}{(1+\theta x)^2} .$$

$$\Rightarrow I_1(\theta) \underset{\substack{\uparrow \\ \text{Info} \\ \text{Equality.}}}{=} -E_\theta\left[-\frac{X^2}{(1+\theta X)^2}\right] = E_\theta\left[\frac{X^2}{(1+\theta X)^2}\right]$$

$$= \int_{-1}^{1} \frac{x^2}{(1+\theta x)^2} \cdot \frac{1}{2}(1+\theta x)\, dx$$

$$= \frac{1}{2}\int_{-1}^{1} \frac{x^2}{1+\theta x}\, dx \quad \longleftarrow \text{use partial fractions here} \quad (*)$$

Write

$$\frac{x^2}{1+\theta x} = \frac{x}{\theta} - \frac{1}{\theta^2} + \frac{1}{\theta^2(1+\theta x)}$$

$$(*) = \frac{1}{2}\int_{-1}^{1} \left[\frac{x}{\theta} - \frac{1}{\theta^2} + \frac{1}{\theta^2(1+\theta x)} \right] dx$$

$$= \frac{1}{2}\left[\frac{x^2}{2\theta} - \frac{x}{\theta^2} + \frac{1}{\theta^2}\log(1+\theta x)\frac{1}{\theta} \right]_{-1}^{1}$$

$$= \frac{1}{2}\left[\cancel{\frac{1}{2\theta}} - \frac{1}{\theta^2} + \frac{\log(1+\theta)}{\theta^3} - \cancel{\frac{1}{2\theta}} - \frac{1}{\theta^2} - \frac{\log(1-\theta)}{\theta^3} \right]$$

$$= \frac{1}{2}\left[-\frac{2}{\theta^2} + \frac{1}{\theta^3}\log\left(\frac{1+\theta}{1-\theta}\right) \right]$$

$$= \frac{1}{2}\left[\frac{-2\theta}{\theta^3} + \frac{1}{\theta^3}\log\left(\frac{1+\theta}{1-\theta}\right) \right]$$

$$= \frac{1}{2}\left[\frac{\log\left(\frac{1+\theta}{1-\theta}\right) - 2\theta}{\theta^3} \right]$$

$$\Rightarrow \sigma^2_{\hat\theta} = \frac{1}{I_1(\theta)} = \frac{2\theta^3}{\log\left(\frac{1+\theta}{1-\theta}\right) - 2\theta},$$

as claimed.

(d) Note that

$$n\,\mathrm{var}_\theta(\tilde\theta) = n\left(\frac{3-\theta^2}{n}\right) = 3-\theta^2.$$

Also

$$\sigma^2_{\hat\theta} = \frac{2\theta^3}{\log\left(\frac{1+\theta}{1-\theta}\right) - 2\theta}$$

$$\overset{\text{hint}}{=} \frac{2\theta^3}{2\left(\theta + \frac{\theta^3}{3} + \frac{\theta^5}{5} + \cdots\right) - 2\theta}$$

$$= \frac{\cancel{2}\theta^3}{\cancel{2}\left(\frac{\theta^3}{3} + \frac{\theta^5}{5} + \cdots\right)}$$

$$= \frac{1}{\frac{\frac{\theta^3}{3} + \frac{\theta^5}{5} + \cdots}{\theta^3}} \xrightarrow[\theta\to 0]{} \frac{1}{(1/3)} = 3.$$

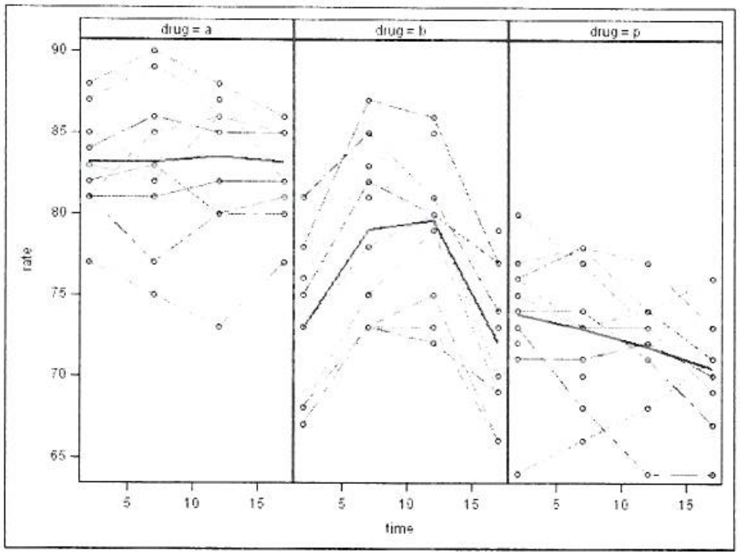
Similarly,

$$\lim_{\theta\to 0} n\,\mathrm{var}_\theta(\tilde\theta) = \lim_{\theta\to 0} (3-\theta^2) = 3.$$

For θ small, $\sigma^2_{\hat\theta}$ & $n\,\mathrm{var}_\theta(\tilde\theta)$ are approximately the same.

Problem 5

(a) The data provided is in SAS format and code for a spaghetti plot is included, producing:



Under the placebo, the mean heartrate starts at about 74 beats per minute (bpm) at two minutes and decreases linearly to about 72 beats per minute at 17 minutes. There is a fair amount of variability in terms of the women, including one who *increases* over the 15 minutes. In terms of the mean, treatment A seems to elevate heartrate to 83 bpm, constant over the 15 minutes. Treatment B seems to initially slightly lower heartrate relative to placebo, then it increases and decreases rather sharply over the 15 minutes. Since the mean trajectories are quite different across treatments, we expect to see a significant time*treatment interaction in the Type III tests.

(b) For $s < t$ we have

$$\begin{aligned} cov(Y_{ijs}, Y_{ijt}) &= cov(\mu + \rho_{i(j)} + \alpha_j + \beta_s + (\alpha\beta)_{js} + \epsilon_{ijs}, \mu + \rho_{i(j)} + \alpha_j + \beta_t + (\alpha\beta)_{jt} + \epsilon_{ijt}) \\ &= cov(\rho_{i(j)} + \epsilon_{ijs}, \rho_{i(j)} + \epsilon_{ijt}) \\ &\overset{ind.}{=} cov(\rho_{i(j)}, \rho_{i(j)}) = var(\rho_{i(j)}) = \sigma_\rho^2. \end{aligned}$$

$$\begin{aligned} cov(Y_{ijs}, Y_{ijs}) &= cov(\mu + \rho_{i(j)} + \alpha_j + \beta_s + (\alpha\beta)_{js} + \epsilon_{ijs}, \mu + \rho_{i(j)} + \alpha_j + \beta_s + (\alpha\beta)_{js} + \epsilon_{ijs}) \\ &= cov(\rho_{i(j)} + \epsilon_{ijs}, \rho_{i(j)} + \epsilon_{ijs}) \\ &= var(\rho_{i(j)} + \epsilon_{ijs}) \overset{ind.}{=} \sigma_\rho^2 + \sigma^2. \end{aligned}$$

So then

$$corr(Y_{ijs}, Y_{ijt}) = \frac{\sigma_\rho^2}{\sigma_\rho^2 + \sigma^2}.$$

(c) The Type III SS table is

Effect	Num DF	Den DF	F Value	Pr > F
drug	2	81	20.25	<.0001
time	3	81	15.56	<.0001
drug*time	6	81	11.20	<.0001

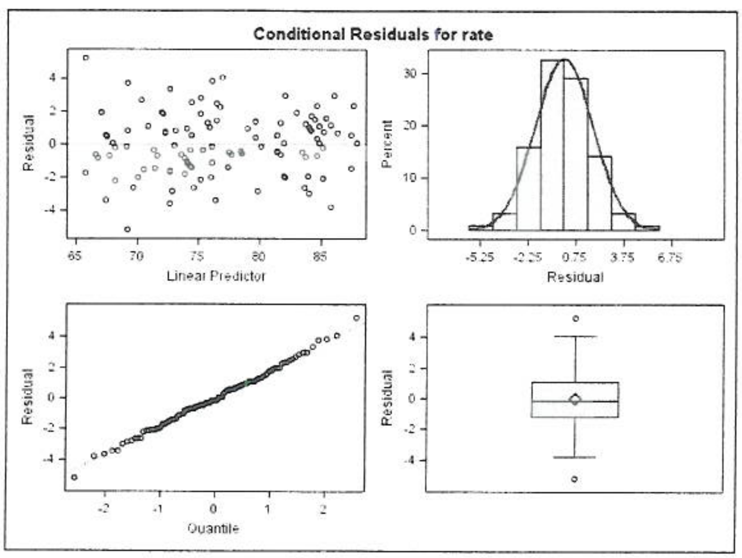
The time*treatment interaction is significant, reflecting that the shape of the mean heartrate profile changes across treatments; this coincides with interaction plots.

(d) Here are the requested pairwise differences, adjusted using Tukey with FER≤ 0.05:

Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha
drug*time	a-b at 2 min	10.0000	1.9301	81	5.18	<.0001	<.0001	0.05
drug*time	a-p at 2 min	9.1000	1.9301	81	4.71	<.0001	<.0001	0.05
drug*time	b-p at 2 min	-0.9000	1.9301	81	-0.47	0.6423	0.6423	0.05
drug*time	a-b at 7 min	4.1000	1.9301	81	2.12	0.0367	0.0367	0.05
drug*time	a-p at 7 min	10.3000	1.9301	81	5.34	<.0001	<.0001	0.05
drug*time	b-p at 7 min	6.2000	1.9301	81	3.21	0.0019	0.0019	0.05
drug*time	a-b at 12 min	3.8000	1.9301	81	1.97	0.0524	0.0524	0.05
drug*time	a-p at 12 min	11.4000	1.9301	81	5.91	<.0001	<.0001	0.05
drug*time	b-p at 12 min	7.6000	1.9301	81	3.94	0.0002	0.0002	0.05
drug*time	a-b at 17 min	10.9000	1.9301	81	5.65	<.0001	<.0001	0.05
drug*time	a-p at 17 min	12.6000	1.9301	81	6.53	<.0001	<.0001	0.05
drug*time	b-p at 17 min	1.7000	1.9301	81	0.88	0.3810	0.3810	0.05

Treatment A is induces significantly larger mean heartrates than B at all four time points. Treatment A induces significantly larger mean heartrates than placebo at all four time points. Treatment B induces significantly larger mean heartrates than placebo at 7 and 12 minutes, but there is no significant difference at 2 and 17 minutes.

(e) The conditional residual plots $e_i = Y_{ijk} - \hat{\mu} - \hat{\rho}_{i(j)} - \hat{\alpha}_j - \hat{\beta}_k - \widehat{(\alpha\beta)}_{jk}$ show constant variance and normality assumptions are okay:



(f) The tests for normality of the $\hat{\rho}_{i(j)}$ from proc univariate show no cause for alarm:

Tests for Normality

Test	--Statistic---		-----p Value------	
Shapiro-Wilk	W	0.950776	Pr < W	0.1773
Kolmogorov-Smirnov	D	0.129771	Pr > D	>0.1500

```
Cramer-von Mises      W-Sq  0.086703    Pr > W-Sq   0.1668
Anderson-Darling      A-Sq  0.547273    Pr > A-Sq   0.1491

                      Normal Probability Plot
    5.5+                                       +++  *
       |                                    *+* *
       |                                 +*+
       |                            *******
       |                          **+++
       |                        **++
       |                      **+
       |                   ***+
       |                 +*+
       |              +++*
       |           +++ **
       |        +*+*
       |     *++
   -7.5+  +++
        +----+----+----+----+----+----+----+----+----+----+
```

#6 The likelihood function (towards finding a sufficient statistic) is

$$L(\theta|\underset{\sim}{x}) = \prod_{i=1}^{n} \frac{\theta}{x_i^2} I(x_i \geq \theta)$$

$$= \frac{\theta^n}{\prod_{i=1}^{n} x_i^2} \prod_{i=1}^{n} I(x_i \geq \theta)$$

$$= \underbrace{\theta^n I(x_{(1)} \geq \theta)}_{g(t|\theta)} \underbrace{\frac{1}{\prod_{i=1}^{n} x_i^2}}_{h(\underset{\sim}{x})}$$

By the Factorization Theorem

$$T = T(\underset{\sim}{X}) = X_{(1)}$$

is a sufficient statistic.

Find distribution of T. The pdf of $T = X_{(1)}$ is

$$f_{X_{(1)}}(x|\theta) \underset{\substack{\uparrow \\ x_{(1)} \geq \theta}}{=} n f_X(x|\theta)\left[1 - F_X(x|\theta)\right]^{n-1},$$

where

$$F_X(x|\theta) = \int_\theta^x \frac{\theta}{s^2}\, ds$$

$$= -\frac{\theta}{s}\Big|_\theta^x = \frac{\theta}{s}\Big|_x^\theta$$

CDF of X $\longrightarrow$ $$= 1 - \frac{\theta}{x}, \quad x \geq \theta.$$

Therefore,

$$f_{X_{(1)}}(x|\theta) = n\left(\frac{\theta}{x^2}\right)\left(\frac{\theta}{x}\right)^{n-1} I(x \geq \theta)$$

$$= \frac{n\theta^n}{x^{n+1}} I(x \geq \theta).$$

Show $T \equiv X_{(1)}$ has MLR.

Suppose $\theta_2 > \theta_1$. It suffices to show that

$$\frac{f_{X_{(1)}}(t|\theta_2)}{f_{X_{(1)}}(t|\theta_1)}$$

is an increasing (a non-decreasing) function of t.

$$\frac{f_{X_{(1)}}(t|\theta_2)}{f_{X_{(1)}}(t|\theta_1)} = \frac{\frac{n\theta_2^n}{t^{n+1}} I(t \geq \theta_2)}{\frac{n\theta_1^n}{t^{n+1}} I(t \geq \theta_1)}$$

$$= \left(\frac{\theta_2}{\theta_1}\right)^n \frac{I(t \geq \theta_2)}{I(t \geq \theta_1)}$$

This is a constant function of t, which is non-decreasing.

Therefore the sufficient statistic T has MLR.

(b) A direct appeal to Karlin Rubin gives the UMP level α test. This test uses the test function

$$\delta(T) = \begin{cases} 1, & T > c \\ 0, & \text{o.w.}, \end{cases}$$

where $T = X_{(1)}$ and where c satisfies

$$\alpha = P_{\theta_0}(T > c).$$

Note that

$$P_{\theta_0}(T > c) = \int_c^\infty \frac{n\theta_0^n}{t^{n+1}}\,dt$$

$$= \left(-\frac{1}{n}\right)\frac{n\,\theta_0^n}{t^n}\Big|_c^\infty$$

$$= \frac{\theta_0^n}{t^n}\Big|_\infty^c = \frac{\theta_0^n}{c^n} - 0$$

Therefore,

$$\alpha = \frac{\theta_0^n}{c^n} \rightarrow c^n = \frac{\theta_0^n}{\alpha}$$

$$\rightarrow c = \frac{\theta_0}{\alpha^{1/n}}.$$

The UMP level rejection region is

$$R = \{ \underset{\sim}{x} : x_{(1)} > \theta_0 \alpha^{-1/n} \}.$$

(c) From part (b), the "acceptance region" is

$$A(\theta_0) = \{ \underset{\sim}{x} : x_{(1)} \le \theta_0 \alpha^{-1/n} \}$$

Define

$$C(\underset{\sim}{x}) = \{ \theta : x_{(1)} \le \theta \alpha^{-1/n} \}$$

Therefore

$$\begin{aligned} 1-\alpha &= P_\theta (\underset{\sim}{X} \in A(\theta)) \\ &= P_\theta (\theta \in C(\underset{\sim}{X})) \\ &= P_\theta (X_{(1)} \le \theta \alpha^{-1/n}) \\ &= P_\theta (\alpha^{1/n} X_{(1)} \le \theta) \end{aligned}$$

Showing that $[\alpha^{1/n} X_{(1)}, \infty)$ is a $100(1-\alpha)\%$ confidence set.

Another approach: One can show that

$$Q \equiv Q(X_{(1)}, \theta) = \frac{\theta}{X_{(1)}} \sim \text{beta}(n, 1)$$

Therefore, Q is a pivot & we can write

$$1-\alpha = P_\theta \Big(\underbrace{b_{n,1,\frac{\alpha}{2}}}_{\text{lower quantile}} \le \frac{\theta}{X_{(1)}} \le \underbrace{b_{n,1,1-\frac{\alpha}{2}}}_{\text{upper quantile}} \Big)$$

and obtain a CI from this.

Ans: $\left(b_{n,1,\frac{\alpha}{2}} X_{(1)},\ b_{n,1,1-\frac{\alpha}{2}} X_{(1)} \right)$