SOLUTION TO BASIC PHD WRITTEN EXAMINATION

[Day 1, Problem 1]

(a) (4 points) $E \log Y = E\left[E\left(\sum_{i=1}^{N} \log X_i | N\right)\right] = \lambda\mu$, and

$$\begin{aligned} \mathrm{var}\left[\log Y\right] &= E\mathrm{var}\left[\sum_{i=1}^{N} \log X_i \middle| N\right] + \mathrm{var}\left(E\left[\sum_{i=1}^{N} \log X_i \middle| N\right]\right) \\ &= E[N\sigma^2] + \mathrm{var}[N\mu] \\ &= \lambda(\sigma^2 + \mu^2). \end{aligned}$$

(b) (5 points)

$$\begin{aligned} EY^t &= E\left[\prod_{i=1}^{N} X_i^t\right] \\ &= E\left[\left(E\left[X_1^t\right]\right)^N\right] \\ &= E\left[M(t)^N\right] \\ &= E\exp\left[N \log M(t)\right] \\ &= \exp\left[\lambda(M(t) - 1)\right], \end{aligned}$$

where the last equality follows from the recognition that the right-hand-side of the second-to-last equality is the moment generating function of N, specifically of the form Ee^{uN} with $u = \log M(t)$.

(c) (7 points) Note that $Y^{1/\lambda} = e^U$, where $U = \lambda^{-1}\sum_{i=1}^{N} \log X_i$. Using conditioning arguments similar to those used before, we have that

$$E\left[\lambda^{-1}\sum_{i=1}^{N}(\log X_i - \mu)\right]^2 = \frac{\sigma^2}{\lambda} \to 0,$$

as $\lambda \to \infty$, and thus $U - N\mu/\lambda \to_p 0$, as $\lambda \to \infty$. Furthermore,

$$E\left[N\mu/\lambda - \mu\right]^2 = E\left[(N - \lambda)\mu/\lambda\right]^2 = \mu^2/\lambda \to 0,$$

as $\lambda \to \infty$, and thus $U \to_p \mu$, as $\lambda \to \infty$. The conclusion now follows by Slutsky's theorem.

(d) (9 points) Let $W = \left(e^{-\lambda\mu} Y\right)^{1/\tau}$, and note that

$$
\begin{aligned}
E\left[e^{t\log W}\right] &= E\left[\exp\left(-\lambda\mu t/\tau + t\tau^{-1}\sum_{i=1}^{N}\log X_i\right)\right] \\
&= \exp\left(-\lambda\mu t/\tau + \lambda\left[M(t\tau^{-1}) - 1\right]\right) \\
&= \exp\left(-\lambda\mu t/\tau + \lambda\left[M(0) - 1 + \dot{M}(0)t\tau^{-1} + \ddot{M}(0)\frac{t^2}{2\tau^2} + o(t^2\tau^{-2})\right]\right) \\
&= \exp\left(-\lambda\mu t/\tau + \lambda\mu t/\tau + \frac{(\sigma^2+\mu^2)t^2}{2(\sigma^2+\mu^2)} + o(t^2\lambda^{-1})\right) \\
&\to e^{t^2/2},
\end{aligned}
$$

as $\lambda \to \infty$, and where $0 \leq t \leq \delta\lambda^{1/2}$ and the second equality follows from part (c). Hence $\log W \to_d N(0,1)$, as $\lambda \to \infty$, and the desired conclusion now follows by another application of Slutsky's theorem.

[**Day 1, Problem 2**]

(a) (4 points) Let $f(x)$ and $g(x)$ be the Radon-Nikodym derivatives of $F(x)$ and $G(x)$ with respect to $F(x) + G(x)$, respectively. Then, the density of X is $\theta f(x) + (1-\theta)g(x)$. For $0 \leq \theta_1 < \theta_2 \leq 1$, we have

$$\frac{\theta_2 f(x) + (1-\theta_2)g(x)}{\theta_1 f(x) + (1-\theta_1)g(x)} = \frac{\theta_2 f(x)/g(x) + (1-\theta_2)}{\theta_1 f(x)/g(x) + (1-\theta_1)}.$$

Therefore, the family of densities of X is MLR in $Y(X) = f(X)/g(X)$. Hence, a UMP test is given by

$$T = \begin{cases} 1 & Y(X) > c, \\ \gamma & Y(X) = c, \\ 0 & Y(X) < c, \end{cases}$$

where c and γ are determined by $E[T(X)] = \alpha$ when $\theta = \theta_0$.

(b) (6 points) For any test T, its power is

$$\begin{aligned} \beta_T(\theta) &= \int T(x)[\theta f(x) + (1-\theta)g(x)]d(F+G) \\ &= \theta \int T(x)[f(x) - g(x)]d(F+G) + \int T(x)g(x)d(F+G), \end{aligned}$$

which is a linear function of θ. Since T has level α and $\beta_T(\theta)$ is a linear function, $\beta_T(\theta) \leq \alpha$ for any $\theta \in [0,1]$. Therefore, $T(x) \equiv \alpha$ is a UMP test.

(c) (6 points) Suppose $T^*(x)$ is a UMP. Let $a = \int T^*(x)[f(x) - g(x)]d(F+G)$. If $a > 0$, $\beta_{T^*}(\theta) < \alpha$ for $\theta < \theta_1$. Therefore $T^*(x)$ is not as powerful as $T(x) \equiv \alpha$. Similarly, if $a < 0$, $T^*(x)$ is also not as powerful as $T(x) \equiv \alpha$. Hence, any test with nonconstant power function cannot be UMP. Next, we prove that $T(x) \equiv \alpha$ is also not UMP. From part (a), we see that the UMP test of size α for testing $H_0 : \theta \leq \theta_2$ versus $H_1 : \theta > \theta_2$ has power $> \alpha$ at $\theta_0 \in (\theta_2, 1]$, i.e. it's more powerful than $T(x) \equiv \alpha$ at θ_0. Hence, $T(x) \equiv \alpha$ cannot be a UMP.

(d) (5 points) Suppose $\tilde{T}(x)$ is an unbiased test. Then, $\beta_{\tilde{T}}(\theta) \leq \alpha$ for $\theta \in [\theta_1, \theta_2]$ and $\beta_{\tilde{T}}(\theta) \geq \alpha$ for $\theta \notin [\theta_1, \theta_2]$. Since the power function is linear, only tests with constant power can be unbiased. Therefore, $T(x) \equiv \alpha$ is UMPU.

(e) (4 points) The likelihood

$$\ell(\theta) = \theta[f(X) - g(X)] + g(X).$$

We have

$$\sup_{0\le\theta\le 1} \ell(\theta) = \begin{cases} f(X) & \text{when } f(X) \ge g(X), \\ g(X) & \text{when } f(X) < g(X). \end{cases}$$

For $0 < \theta_1 \le \theta_2 < 1$,

$$\sup_{0<\theta_1\le\theta_2<1} \ell(\theta) = \begin{cases} \theta_2[f(X) - g(X)] + g(X), & \text{when } f(X) \ge g(X), \\ \theta_1[f(X) - g(X)] + g(X), & \text{when } f(X) < g(X). \end{cases}$$

Therefore, the likelihood ratio test statistic

$$\lambda(X) = \begin{cases} \frac{\theta_2[f(X)-g(X)]+g(X)}{f(X)}, & \text{when } f(X) \ge g(X), \\ \frac{\theta_1[f(X)-g(X)]+g(X)}{g(X)}, & \text{when } f(X) < g(X). \end{cases}$$

[**Day 2, Problem 1**]

(a) (2 points)

(b) (3 points)

(c) (4 points)

(d) (5 points)

(e) (5 points)

(f) (6 points)

[**Day 2, Problem 2**]

(a) (1 point) X would have to be linear in Y, i.e. supported on two points, which is impossible since $X \sim$ normal.

(b) (6 points) $\rho = \text{cov}(X,Y)/\sqrt{\theta(1-\theta)}$. Since θ is fixed, to maximize ρ we need to maximize $\text{cov}(X,Y) = \text{E}[XY] = \text{E}[XY|Y=0]P(Y=0) + \text{E}[XY|Y=1]P(Y=1) = \text{E}[XY|Y=1]P(Y=1) = \text{E}[X|Y=1]\theta$. Hence we need to maximize $\text{E}[X|Y=1]$

In the xy-plane, the joint pdf of (X,Y) is concentrated on the lines $y=0$ and $y=1$. Hence we maximize $\text{E}[X|Y=1]$ by putting all the (positive) density on $x > c$ if $y=1$ and on $x < c$ if $y=0$, for an appropriate c. i.e. we put all the density on $\{(x,y) : x < c, y = 0)\} \cup \{(x,y) : x > c, y = 1)\}$. i.e. $Y = I(X > c)$ w.p. 1. Of course, $c = \Phi^{-1}(1-\theta)$ to guarantee that $Y \sim \text{Bernoulli}(\theta)$.

(c) (6 points) Now we compute $\text{cov}(X,Y) = \text{E}[XY] = \theta\text{E}[X|Y=1]$:

$\text{E}[X|Y=1] = \text{E}[X|X>c] = \int_c^\infty x\phi(x)dx/\{1-\Phi(c)\}$
$= \phi(c)/\{1-\Phi(c)\} = \phi(\Phi^{-1}(1-\theta))/\theta$,
using $d\phi(x) = -x\phi(x)dx$ in the integration.

Then

$$\text{cov}(X,Y) = \theta\text{E}[X|Y=1] = \phi(\Phi^{-1}(1-\theta)),$$

and

$$\rho^* = \text{corr}(X,Y) = \frac{\phi(\Phi^{-1}(1-\theta))}{\sqrt{\theta(1-\theta)}}$$

For $\theta = 0.001$: $c = 3.09, \rho^* = 0.107$.

(d) (6 points) View ρ^* as a function of c and show that the derivative is negative for $c > 0$ and positive for $c < 0$, so the mode is at $c = 0$, i.e. $\theta = 0.5$. That follows by observing that $d\phi(x)$ is negative for positive x, and vice versa.

So $\rho^{**} = \phi(0)/0.5 = \sqrt{2/\pi} \approx 0.7979$.

(e) (3 points) Define $T_i = X_iY_i$ then using sample means

$$\hat{\rho} = \frac{\bar{T}}{\sqrt{\bar{Y}(1-\bar{Y})}}.$$

(f) (3 points) To get the asymptotic variance, apply the delta method to the vector $(\bar{T}, \bar{Y})^\top$. The sample covariance matrix can be used to estimate the covariance matrix of $(T_i, Y_i)^\top$.

[This can be done as a GEE as well. Same final answer]

Important: MLE is clearly not possible, the stated model does not fully determine the pdf of (X_i, Y_i). Any attempt at MLE deserves 0 credit.