Homework 0

1 (a) **Required To Prove:** $\mathrm{E}[\mathrm{E}[X|Y]] = \mathrm{E}[X]$.

Proof:

$$\begin{aligned}
E[E[X|Y]] &= \int\limits_y E[X|Y]p(Y=y)dy \\
&= \int\limits_y (\int\limits_x x \cdot p(X=x|Y=y)dx)\ p(Y=y)dy \\
&= \int\limits_y (\int\limits_x x \cdot \frac{p(X=x,Y=y)}{p(Y=y)}dx)\ p(Y=y)dy \\
&= \int\limits_y \int\limits_x x \cdot p(X=x,Y=y)dxdy \\
&= \int\limits_x x \int\limits_y p(X=x,Y=y)dydx \\
&= \int\limits_x x \cdot p(X=x)\ dx \\
&= E[X]
\end{aligned}$$

$\square$

(b) **Required To Prove:** $\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathrm{E}[X|Y])$.

Proof:

$$\begin{aligned}
E[Var(X|Y)] + Var(E[X|Y]) &= \int\limits_y E[E[X^2|Y] - E[X|Y]^2] + E[E[X|Y]^2] - E[E[X|Y]]^2 dy \\
&= \int\limits_y E[E[X^2|Y]] - E[E[X|Y]]^2 dy \qquad (1)
\end{aligned}$$

From the proof of problem (a), we get,

$$\begin{aligned}
E[E[X^2|Y]] &= E[X^2] \\
E[E[X|Y]]^2 &= E[X]^2
\end{aligned}$$

Thus, we can rewrite Eq.(1) as,

$$\begin{aligned}
E[Var(X|Y)] + Var(E[X|Y]) &= E[X^2] - E[X]^2 \\
&= Var(X)
\end{aligned}$$

$\square$

2. We can obtain the density function $p(y)$ using two approaches:

Approach 1: a. , b.

First, $Y = f(X) = \sqrt(X)$. This implies,

$$X = f^{-1}(Y) = Y^2 \tag{2}$$

$$\begin{aligned} p(y) &= \left|\frac{d}{dy} f^{-1}(y)\right| p(f^{-1}(y)) \\ \implies p(y) &= \left|\frac{d}{dy}(y^2)\right| p(y^2) \qquad \text{(from 2)} \\ \implies p(y) &= 2y\lambda e^{-\lambda y^2} \end{aligned}$$

The cdf is defined as follows:

$$\begin{aligned} F(y) &= \int\limits_{-\infty}^{y} p(y)dy \\ \implies F(y) &= \lambda \int\limits_{-\infty}^{y} 2ye^{-\lambda y^2} dy \\ \implies F(y) &= 1 - e^{-\lambda y^2} \qquad \text{(variable substitution)} \end{aligned} \tag{3}$$

We can verify from eq. 3 that $F(0) = 0$ and $F(\infty) = 1$.

Approach 2: a. , b.

Consider the cdf F:

$$\begin{aligned} F(y) &= p(Y \le y) = p(\sqrt{X} \le y) \\ \implies F(y) &= p(X \le y^2) = \int\limits_{0}^{y^2} \lambda e^{-\lambda x} dx \\ \implies F(y) &= 1 - e^{-\lambda y^2} \end{aligned} \tag{4}$$

We can verify from eq. 4 that $F(0) = 0$ and $F(\infty) = 1$.

Differentiating $F(y)$ wrt y, we get the density function as :

$$p(y) = 2y\lambda e^{-\lambda y^2}$$

c.

$$\begin{aligned} F(y) &= 1 - e^{-\lambda y^2} \\ \implies y &= \sqrt{\frac{-ln(1-F)}{\lambda}} \\ \implies F^{-1}(y) &= \sqrt{\frac{-ln(1-F)}{\lambda}} \end{aligned}$$

d.

Here we use *integration by parts* which is stated as follows:

$$\int u dv = uv - \int v du \tag{5}$$

Mean

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y \cdot p(y) dy = 2\int_{0}^{\infty} y \cdot p(y) dy \\ &= 2\int_{0}^{\infty} y \cdot (2\lambda \exp^{-\lambda y^2} y) dy \\ &= 2\int_{0}^{\infty} -y \cdot d(\exp^{-\lambda y^2}) \\ &= 2[-y \cdot \exp^{-\lambda y^2}]_0^\infty + 2\int_{0}^{\infty} \exp^{-\lambda y^2} dy \qquad \text{(from Eq. 5)} \\ &= \sqrt{\frac{\pi}{\lambda}} \end{aligned}$$

Variance

$$\begin{aligned} Var(Y) &= E(Y^2) - E(Y)^2 \\ &= \int_{-\infty}^{\infty} y^2 \cdot p(y) dy - \frac{\pi}{\lambda} \\ &= 2\int_{0}^{\infty} y^2 \cdot (2\lambda \exp^{-\lambda y^2} y) dy - \frac{\pi}{\lambda} \\ &= 2\int_{0}^{\infty} \lambda y^2 \cdot (\exp^{-\lambda y^2}) d(y^2) - \frac{\pi}{\lambda} \qquad \text{(var substitution)} \end{aligned}$$

Changing variable as $y^2 = t$ ($y = 0 \implies t = 0, y = \infty \implies t = \infty$), we get,

$$\begin{aligned} Var(Y) &= 2\int_{0}^{\infty} \lambda t \cdot (\exp^{-\lambda t}) d(t) - \frac{\pi}{\lambda} \\ &= 2\int_{0}^{\infty} -t \cdot d(\exp^{-\lambda t}) - \frac{\pi}{\lambda} \\ &= 2[-t \cdot \exp^{-\lambda t}]_0^\infty + 2\int_{0}^{\infty} \exp^{-\lambda t} dt - \frac{\pi}{\lambda} \qquad \text{(from Eq. 5)} \\ &= 0 + 2\int_{0}^{\infty} \exp^{-\lambda t} dt - \frac{\pi}{\lambda} \\ &= 2[-\frac{1}{\lambda} \cdot \exp^{-\lambda t}]_0^\infty - \frac{\pi}{\lambda} \\ &= \frac{2-\pi}{\lambda} \end{aligned}$$

3. The maximum likelihood estimate for λ can be obtained by maximizing the joint log-likelihood

$$\begin{aligned} L(\lambda; y_1, y_2, \cdots y_n) &= p(y_1, y_2, \cdots y_n; \lambda) \\ &= \prod_{i=1}^{N} p(y_i; \lambda) \\ &= \prod_{i=1}^{N} 2\lambda y_i e^{-\lambda y_i^2} \end{aligned} \tag{6}$$

Differentiating logarithm of eq. (6) w.r.t. λ, we get,

$$
\begin{aligned}
l = ln\ L(\lambda; y_1, y_2, \cdots y_n) &= \sum_{i=1}^{N} ln\ 2y_i \lambda e^{-\lambda y_i^2} \\
\implies \frac{dl}{d\lambda} &= \frac{N}{\lambda} - \sum_{i=1}^{N} y_i^2 = 0 \\
\implies \hat{\lambda} &= \frac{N}{\sum_{i=1}^{N} y_i^2}
\end{aligned}
$$

So the maximum likelihood estimate for λ is $\frac{N}{\sum_{i=1}^{N} y_i^2}$