MATH 556 - ASSIGNMENT 3: SOLUTIONS

1 (a) Suppose first that

$$\mathbb{E}_{f_X}[g(X)] \geq \mathbb{E}_{f_Y}[g(Y)]$$

for any non-decreasing real function g. Proof given for continuous random variables, but proof in the discrete case follows after minor adjustment. Let $g_t(x) = \mathbb{I}_{[t,\infty)}(x)$; this is a non-decreasing function for all real t, so

$$\mathbb{E}_{f_X}[g_t(X)] \geq \mathbb{E}_{f_Y}[g_t(Y)]$$

But

$$\mathbb{E}_{f_X}[g_t(X)] = \int_{-\infty}^{\infty} \mathbb{I}_{[t,\infty)}(x) f_X(x)\,dx = \int_t^{\infty} f_X(x)\,dx = P[X > t]$$

and hence for all t,

$$P[X > t] \geq P[Y > t] \qquad \Longleftrightarrow \qquad F_X(t) \leq F_Y(t)$$

by linearity of expectation.

Conversely, suppose X is stochastically greater than Y, so that for all t,

$$F_X(t) \leq F_Y(t) \qquad \Longleftrightarrow \qquad P[X \geq Y] = 1$$

Hence, for any non-decreasing g

$$P[g(X) \geq g(Y)] \geq P[X \geq Y] = 1$$

and thus if $Z = g(X) - g(Y)$ then

$$P[Z \geq 0] = P[g(X) - g(Y) \geq 0] = 1.$$

Hence, as Z is a non-negative random variable,

$$\mathbb{E}[Z] = \mathbb{E}[g(X) - g(Y)] \geq 0 \qquad \Longrightarrow \qquad \mathbb{E}_{f_X}[g(X)] \geq \mathbb{E}_{f_Y}[g(Y)]$$

6 MARKS

(b) Three different ways to prove this:

- Suppose first that $Z(n) \sim Gamma(n, \lambda)$, for positive integer n. Then, integrating by parts,

$$\begin{aligned}
P[Z(n) > 1] &= \int_1^{\infty} \frac{\lambda^n}{\Gamma(n)} z^{n-1} e^{-\lambda x}\,dz \\
&= \left[-\frac{\lambda^{n-1}}{\Gamma(n)} z^{n-1} e^{-\lambda z} \right]_1^{\infty} + \int_1^{\infty} \frac{\lambda^{n-1}}{\Gamma(n-1)} z^{n-2} e^{-\lambda z}\,dz \\
&= \frac{\lambda^{n-1}}{\Gamma(n)} e^{-\lambda} + \int_1^{\infty} \frac{\lambda^{n-1}}{\Gamma(n-1)} z^{n-2} e^{-\lambda z}\,dz \\
&= \frac{\lambda^{n-1}}{\Gamma(n)} e^{-\lambda} + \left[-\frac{\lambda^{n-2}}{\Gamma(n)} z^{n-2} e^{-\lambda z} \right]_1^{\infty} + \int_1^{\infty} \frac{\lambda^{n-2}}{\Gamma(n-1)} z^{n-3} e^{-\lambda z}\,dz \\
&= \frac{\lambda^{n-1}}{\Gamma(n)} e^{-\lambda} + \frac{\lambda^{n-2}}{\Gamma(n-1)} e^{-\lambda} + \int_1^{\infty} \frac{\lambda^{n-2}}{\Gamma(n-1)} z^{n-3} e^{-\lambda z}\,dz
\end{aligned}$$

Iterating this procedure, we ultimately obtain

$$\begin{aligned}
P[Z(n) > 1] &= \sum_{j=1}^{n-1} \frac{\lambda^{n-j}}{\Gamma(n-j+1)} e^{-\lambda} + \int_1^\infty \frac{\lambda}{\Gamma(2)} e^{-\lambda z}\, dz \\
&= \sum_{j=1}^{n-1} \frac{\lambda^{n-j}}{\Gamma(n-j+1)} e^{-\lambda} + \left[-\frac{1}{\Gamma(2)} e^{-\lambda z} \right]_1^\infty \\
&= \sum_{j=1}^{n-1} \frac{\lambda^{n-j}}{\Gamma(n-j+1)} e^{-\lambda} + e^{-\lambda} \\
&= \sum_{j=1}^{n} \frac{\lambda^{n-j}}{\Gamma(n-j+1)} e^{-\lambda} \equiv \sum_{x=0}^{n-1} \frac{\lambda^{x}}{\Gamma(x+1)} e^{-\lambda}
\end{aligned}$$

But $\Gamma(x+1) = x!$ if x is a positive integer, so

$$P[Z(n) > 1] = \sum_{x=0}^{n-1} \frac{\lambda^x}{x!} e^{-\lambda} = P[V \leq n-1] = F_V(n-1)$$

where $V \sim Poisson(\lambda)$. This is a fundamental representation of the cdf of the Poisson distribution.

Thus we must have, in the original problem, for real t,

$$F_X(t) = P[X \leq t] = P[Z(t+1) > 1]$$

where $Z(t+1) \sim Gamma(t+1, \lambda)$. Thus,

$$\begin{aligned}
F_X(t) = P[X \leq t] &= \sum_{x=0}^{t} \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)} \\
&\equiv \int_1^\infty \frac{\lambda^{t+1}}{\Gamma(t+1)} z^t e^{-\lambda x}\, dz \\
&= \int_{\lambda/\mu}^\infty \frac{\lambda^{t+1}}{\Gamma(t+1)} \left(\frac{\mu y}{\lambda}\right)^t e^{-\mu y} \left(\frac{\mu}{\lambda}\right) dy \qquad \text{setting } y = \frac{\lambda}{\mu} z \\
&= \int_{\lambda/\mu}^\infty \frac{\mu^{t+1}}{\Gamma(t+1)} y^t e^{-\mu y}\, dy \\
&\leq \int_1^\infty \frac{\mu^{t+1}}{\Gamma(t+1)} y^t e^{-\mu y}\, dy \qquad \text{as } \frac{\lambda}{\mu} \geq 1 \\
&= P[Y \leq t] \\
&= F_Y(t)
\end{aligned}$$

and the result follows.

This result uses the connection between the Poisson and Gamma distributions through the *Poisson Process.*

- By properties of the Poisson distribution, we may write

$$X \stackrel{d}{=} X_1 + X_2$$

where $X_1 \sim Poisson(\mu)$ and $X_2 \sim Poisson(\lambda - \mu)$ are independent random variables, that is X and $X_1 + X_2$ have the same distribution. But, by construction, $X_1 \stackrel{d}{=} Y$, so also

$$X \stackrel{d}{=} Y + X_2$$

and thus for real t, as $P[X_2 \geq 0] = 1$,

$$F_X(t) = P[X \leq t] \equiv P[Y + X_2 \leq t] \leq P[Y \leq t] = F_Y(t)$$

- Suppose $X \sim Poisson(\lambda)$, and suppose $Y|X = x \sim Binomial(x, \theta)$, for some $0 < \theta < 1$, and $x \in \{0, 1, \ldots\}$. Then clearly $P[X \geq Y] = 1$, but also, for $y \in \{0, 1, \ldots\}$,

$$\begin{aligned} f_Y(y) = \sum_{x=0}^{y} f_{Y|X}(y|x) f_X(x) &= \sum_{x=y}^{\infty} \binom{x}{y} \theta^y (1-\theta)^{x-y} \frac{\lambda^x}{x!} e^{-\lambda} \\ &= (\theta\lambda)^y \frac{e^{-\lambda}}{y!} \sum_{x=y}^{\infty} \frac{((1-\theta)\lambda)^{x-y}}{(x-y)!} \\ &= (\theta\lambda)^y \frac{e^{-\lambda}}{y!} \sum_{x=0}^{\infty} \frac{((1-\theta)\lambda)^{x}}{(x)!} \\ &= (\theta\lambda)^y \frac{e^{-\lambda}}{y!} e^{(1-\lambda)\theta} = (\theta\lambda)^y \frac{e^{-\theta\lambda}}{y!} \end{aligned}$$

so $Y \sim Poisson(\theta\lambda)$. Setting $\theta = \mu/\lambda$ gives the result.

4 MARKS

2 (a) Clearly $\mathbb{H}(f_1, f_2) \geq 0$ as the integrand is positive. Also,

$$\begin{aligned} 2\mathbb{H}(f_1, f_2) &= \int_{-\infty}^{\infty} \left(\sqrt{f_1(x)} - \sqrt{f_2(x)}\right)^2 \, dx \\ &= \int_{-\infty}^{\infty} f_1(x)\, dx - 2\int_{-\infty}^{\infty} \sqrt{f_1(x) f_2(x)}\, dx + \int_{-\infty}^{\infty} f_2(x)\, dx \leq 2 \end{aligned}$$

as the first and third terms are 1, and the middle integral is non-negative.

2 MARKS

(b) First note $\mathbb{H}(f_1, f_2) = \mathbb{H}(f_2, f_1)$. Secondly, from (a)

$$\begin{aligned} 2\mathbb{H}(f_1, f_2) &= \int_{-\infty}^{\infty} \left(\sqrt{f_1(x)} - \sqrt{f_2(x)}\right)^2 \, dx \\ &= 2 - 2\int_{-\infty}^{\infty} \sqrt{f_1(x) f_2(x)}\, dx \\ &= 2 - 2\int_{-\infty}^{\infty} \sqrt{\frac{f_2(x)}{f_1(x)}} f_1(x)\, dx \end{aligned}$$

Now, as given in the proof of the Kullback-Leibler non-negativity result, $\log(x+1) \leq x$, so

$$\log \frac{f_2(x)}{f_1(x)} = 2 \log \left(1 + \left(\sqrt{\frac{f_2(x)}{f_1(x)}} - 1 \right) \right) \leq 2 \left(\sqrt{\frac{f_2(x)}{f_1(x)}} - 1 \right)$$

so

$$\int_{-\infty}^{\infty} \log \frac{f_2(x)}{f_1(x)} f_1(x) \, dx \leq \int_{-\infty}^{\infty} 2 \left(\sqrt{\frac{f_2(x)}{f_1(x)}} - 1 \right) f_1(x) \, dx$$

so that

$$\begin{aligned} \mathbb{K}(f_1, f_2) = -\int_{-\infty}^{\infty} \log \frac{f_2(x)}{f_1(x)} f_1(x) \, dx &\geq \int_{-\infty}^{\infty} 2 \left(1 - \sqrt{\frac{f_2(x)}{f_1(x)}} \right) f_1(x) \, dx \\ &= 2 - 2 \int_{-\infty}^{\infty} \sqrt{f_1(x) f_2(x)} \, dx \\ &= \mathbb{H}(f_1, f_2). \end{aligned}$$

and thus

$$\mathbb{H}(f_1, f_2) \leq \mathbb{K}(f_1, f_2).$$

Clearly, we must have $\mathbb{H}(f_2, f_1) \leq \mathbb{K}(f_2, f_1)$, but as $\mathbb{H}(f_1, f_2) = \mathbb{H}(f_2, f_1)$, we have

$$\mathbb{H}(f_1, f_2) \leq \mathbb{K}(f_1, f_2) \qquad \text{and} \qquad \mathbb{H}(f_1, f_2) \leq \mathbb{K}(f_2, f_1)$$

so

$$\mathbb{H}(f_1, f_2) \leq \min \{ \mathbb{K}(f_1, f_2), \mathbb{K}(f_2, f_1) \}.$$

4 MARKS

$\mathbb{H}(f_1, f_2)$ *is termed the (squared) Hellinger distance.*

3 This result simply uses Jensen's Inequality with the convex function

$$g(x) = xe^x$$

which is convex on the positive real half-line. We have

$$g(\mu) = \mu e^{\mu} \leq \mathbb{E}_{f_X}[g(X)] = \sum_{i=1}^{\infty} p_i x_i e^{x_i}$$

Note that the right hand side need not be finite for the result to hold.

4 MARKS