**Problem 1**

Show the following:

(a) $\mathbb{E}(aY + bZ \mid X) = a\mathbb{E}(Y \mid X) + b\mathbb{E}(Z \mid X)$ for $a, b \in \mathbb{R}$,

(b) $\mathbb{E}(Y \mid X) \geq 0$ if $Y \geq 0$,

(c) $\mathbb{E}(1 \mid X) = 1$,

(d) if $X$ and $Y$ are independent then $\mathbb{E}(Y \mid X) = \mathbb{E}(Y)$,

(e) ('pull-through property') $\mathbb{E}(Yg(X) \mid X) = g(X)\mathbb{E}(Y \mid X)$ for any suitable function $g$,

(f) ('tower property') $\mathbb{E}\{\mathbb{E}(Y \mid X, Z) \mid X\} = \mathbb{E}(Y \mid X) = \mathbb{E}\{\mathbb{E}(Y \mid X) \mid X, Z\}$.

**(a)** By definition, for discrete random variables,

$$\mathbb{E}(aY + bZ \mid X = x) = \sum_{y,z} \Big[ ay + bz \Big] \mathbb{P}(Y = y, Z = z \mid X = x).$$

By the distributive property of summation, we have:

$$\mathbb{E}(aY + bZ \mid X = x) = a\sum_{y,z} y\,\mathbb{P}(Y = y, Z = z \mid X = x) + b\sum_{y,z} z\,\mathbb{P}(Y = y, Z = z \mid X = x).$$

Notice that for fixed $y$,

$$\sum_{z} \mathbb{P}(Y = y, Z = z \mid X = x) = \mathbb{P}(Y = y \mid X = x),$$

and similarly,

$$\sum_{y} \mathbb{P}(Y = y, Z = z \mid X = x) = \mathbb{P}(Z = z \mid X = x).$$

Thus, we obtain:

$$\mathbb{E}(aY + bZ \mid X = x) = a\sum_{y} y\,\mathbb{P}(Y = y \mid X = x) + b\sum_{z} z\,\mathbb{P}(Z = z \mid X = x)$$

$$= a\,\mathbb{E}(Y \mid X = x) + b\,\mathbb{E}(Z \mid X = x).$$

Since this holds for every $x$, we have proven that

$$\mathbb{E}(aY + bZ \mid X) = a\,\mathbb{E}(Y \mid X) + b\,\mathbb{E}(Z \mid X).$$

**(b)**

If $Y \geq 0$, then for any fixed $x$ each term in

$$\mathbb{E}(Y \mid X = x) = \sum_{y} y\,\mathbb{P}(Y = y \mid X = x)$$

1

is nonnegative, so
$$\mathbb{E}(Y \mid X) \geq 0.$$

**(c)**

Setting $Y \equiv 1$,

$$\mathbb{E}(1 \mid X = x) = \sum_{y,z} 1 \cdot \mathbb{P}(Y = y, Z = z \mid X = x) = \sum_{y,z} \mathbb{P}(Y = y, Z = z \mid X = x) = 1,$$

since the sum of the conditional probabilities equals 1. Thus,

$$\mathbb{E}(1 \mid X) = 1.$$

**(d)**

If $X$ and $Y$ are independent, then for all $x$ and $y$

$$\mathbb{P}(Y = y \mid X = x) = \mathbb{P}(Y = y).$$

Thus,
$$\mathbb{E}(Y \mid X = x) = \sum_y y \, \mathbb{P}(Y = y \mid X = x) = \sum_y y \, \mathbb{P}(Y = y) = \mathbb{E}(Y).$$

Hence,
$$\mathbb{E}(Y \mid X) = \mathbb{E}(Y).$$

**(e) Pull-Through Property:**

Let $g$ be any function such that $g(X)$ is $\sigma(X)$-measurable. Then for any fixed $x$,

$$\mathbb{E}\big(Y \, g(X) \mid X = x\big) = \sum_{y,z} y \, g(x) \, \mathbb{P}(Y = y, Z = z \mid X = x)$$

$$= g(x) \sum_{y,z} y \, \mathbb{P}(Y = y, Z = z \mid X = x)$$

$$= g(x) \, \mathbb{E}(Y \mid X = x).$$

Thus,
$$\mathbb{E}(Y \, g(X) \mid X) = g(X) \, \mathbb{E}(Y \mid X).$$

**(f) Tower Property:**

$$\mathbb{E}\{\mathbb{E}(Y|X, Z)|X = x\} = \sum_z \left\{ \sum_y y\mathbb{P}(Y = y|X = x, Z = z)\mathbb{P}(X = x, Z = z|X = x) \right\}$$

$$= \sum_{z} \sum_{y} y \frac{\mathbb{P}(Y = y, X = x, Z = z)}{\mathbb{P}(X = x, Z = z)} \cdot \frac{\mathbb{P}(X = x, Z = z)}{\mathbb{P}(X = x)}$$

$$= \sum_{y} y \mathbb{P}(Y = y | X = x)$$

$$= \mathbb{E}\{\mathbb{E}(Y | X) | X = x, Z = z\}$$

## Problem 2

Conditional variance formula.

How should we define $\operatorname{var}(Y \mid X)$, the conditional variance of $Y$ given $X$?

Show that $\operatorname{var}(Y) = \mathbb{E}(\operatorname{var}(Y \mid X)) + \operatorname{var}(\mathbb{E}(Y \mid X))$.

*Proof.* By definition, the variance of $Y$ is

$$\operatorname{Var}(Y \mid X) = \mathbb{E}\left[\left(Y - \mathbb{E}(Y \mid X)\right)^2 \mid X\right].$$

$$
\begin{aligned}
\operatorname{Var}(Y) &= \mathbb{E}\left[\left(Y - \mathbb{E}(Y)\right)^2\right] \\
&= \mathbb{E}\left[\left(Y - \mathbb{E}(Y \mid X) + \mathbb{E}(Y \mid X) - \mathbb{E}(Y)\right)^2\right] \\
&= \mathbb{E}\left[\left(Y - \mathbb{E}(Y \mid X)\right)^2\right] + 2\,\mathbb{E}\left[\left(Y - \mathbb{E}(Y \mid X)\right)\left(\mathbb{E}(Y \mid X) - \mathbb{E}(Y)\right)\right] \\
&\quad + \mathbb{E}\left[\left(\mathbb{E}(Y \mid X) - \mathbb{E}(Y)\right)^2\right] \\
&= \mathbb{E}\left[\operatorname{Var}(Y \mid X)\right] + \operatorname{Var}\left(\mathbb{E}(Y \mid X)\right) \text{ (The cross products are zeros)}
\end{aligned}
$$

$\square$

## Problem 3

(1) Let $X$ and $Y$ be independent discrete random variables, and let $g, h : \mathbb{R} \to \mathbb{R}$. Show that $g(X)$ and $h(Y)$ are independent.

(2) Show that two discrete random variables $X$ and $Y$ are independent if and only if $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ for all $x, y \in \mathbb{R}$.

(3) More generally, show that $X$ and $Y$ are independent if and only if $f_{X,Y}(x, y)$ can be factorized as the product $g(x)h(y)$ of a function of $x$ alone and a function of $y$ alone.

*Proof.* Since $X$ and $Y$ are independent discrete random variables, we have for any subsets $A, B \subset \mathbb{R}$

$$\mathbb{P}(g(X) = a, h(Y) = b) = \sum_{x,y:g(x)=a,h(y)=b} \mathbb{P}(X = x, Y = y)$$

$$= \sum_{x,y:g(x)=a,h(y)=b} \mathbb{P}(X=x)\mathbb{P}(Y=y)$$

$$= \sum_{x:g(x)=a} \mathbb{P}(X=x) \sum_{y:h(y)=b} \mathbb{P}(Y=y)$$

$$= \mathbb{P}(g(X)=a)\mathbb{P}(h(Y)=b).$$

Thus, $g(X)$ and $h(Y)$ are independent.

(2)

$(\Rightarrow)$Suppose that $X$ and $Y$ are independent.

By the definition of independence, for any sets $A, B \subset \mathbb{R}$, we have:

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

In particular, for singleton sets $\{x\}$ and $\{y\}$, we have:

$$P(X=x, Y=y) = P(X=x)P(Y=y)$$

That is, $f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall x, y \in \mathbb{R}$

$(\Leftarrow)$Suppose that $f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall x, y \in \mathbb{R}$

For any sets $A, B \subset \mathbb{R}$, we have:

$$P(X \in A, Y \in B) = \sum_{x \in A, y \in B} f_{X,Y}(x,y)$$

$$= \sum_{x \in A, y \in B} f_X(x)f_Y(y)$$

$$= \sum_{x \in A} f_X(x) \sum_{y \in B} f_Y(y)$$

$$= P(X \in A)P(Y \in B)$$

Therefore, $X$ and $Y$ are independent. $\square$

(3)

$(\Rightarrow)$Suppose that $X$ and $Y$ are independent.

Then by definition, $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all $x, y \in \mathbb{R}$.

Taking $g(x) = f_X(x)$ and $h(y) = f_Y(y)$, we have the desired factorization.

$(\Leftarrow)$Suppose that $f_{X,Y}(x,y) = g(x)h(y)$ for some functions $g$ and $h$.

We need to show that $X$ and $Y$ are independent.

First, let's compute the marginal PMFs:

$$f_X(x) = \sum_y f_{X,Y}(x,y) = \sum_y g(x)h(y) = g(x)\sum_y h(y)$$

Similarly, $f_Y(y) = \sum_x f_{X,Y}(x,y) = \sum_x g(x)h(y) = h(y)\sum_x g(x)$

Let $C_1 = \sum_y h(y)$ and $C_2 = \sum_x g(x)$.

Note that $C_1$ and $C_2$ are constants, and since $f_{X,Y}$ is a valid PMF, we know

$$\sum_x \sum_y f_{X,Y}(x,y) = \sum_x \sum_y g(x)h(y) = \sum_x g(x) \sum_y h(y) = C_1 C_2 = 1$$

So we have $f_X(x) = g(x)C_1$ and $f_Y(y) = h(y)C_2$

Therefore:

$$f_X(x)f_Y(y) = g(x)C_1 \cdot h(y)C_2 = g(x)h(y)C_1 C_2 = g(x)h(y) = f_{X,Y}(x,y)$$

We've shown that $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, which means $X$ and $Y$ are independent.   $\square$

$\square$

## Problem 4

**Stein's identity.** Let $X$ be $N(\mu, \sigma^2)$.
Show that $\mathbb{E}\{(X - \mu)g(X)\} = \sigma^2 \mathbb{E}(g'(X))$ when both sides exist.

*Proof.* The probability density function of $X$ is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Note that

$$f'_X(x) = -\frac{x-\mu}{\sigma^2} f_X(x),$$

so that

$$(x - \mu)f_X(x) = -\sigma^2 f'_X(x).$$

Thus, we have

$$\mathbb{E}[(X - \mu)g(X)] = \int_{-\infty}^{\infty} (x - \mu)g(x)f_X(x)\, dx$$
$$= -\sigma^2 \int_{-\infty}^{\infty} g(x)f'_X(x)\, dx.$$

Using integration by parts with

$$u = g(x), \quad dv = f'_X(x)\, dx, \quad du = g'(x)\, dx, \quad v = f_X(x),$$

we obtain

$$\int_{-\infty}^{\infty} g(x)f'_X(x)\, dx = \Big[g(x)f_X(x)\Big]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} g'(x)f_X(x)\, dx$$
$$= -\int_{-\infty}^{\infty} g'(x)f_X(x)\, dx \quad \text{(boundary term vanishes)}$$
$$= -\mathbb{E}[g'(X)].$$

Therefore,

$$\mathbb{E}[(X - \mu)g(X)] = -\sigma^2\left[-\mathbb{E}[g'(X)]\right]$$
$$= \sigma^2\,\mathbb{E}[g'(X)].$$

$\square$

## Problem 5

(a) Show that $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$, and deduce that

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty,$$

is a density function if $\sigma > 0$.

(b) Calculate the mean and variance of a standard normal variable.

(c) Show that the $N(0,1)$ distribution function $\Phi$ satisfies

$$(x^{-1} - x^{-3})e^{-\frac{1}{2}x^2} < \sqrt{2\pi}[1 - \Phi(x)] < x^{-1}e^{-\frac{1}{2}x^2}, \quad x > 0.$$

These bounds are of interest because $\Phi$ has no closed form.

(d) Let $X$ be $N(0,1)$, and $a > 0$. Show that $\mathbb{P}(X > x + a/x \mid X > x) \to e^{-a}$ as $x \to 0$.

*Proof.*

(a) To show that $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$ and deduce that $f(x)$ is a density function:

First, let's consider the square of the integral:

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2}\,dx\right)^2 = \int_{-\infty}^{\infty} e^{-x^2}\,dx\int_{-\infty}^{\infty} e^{-y^2}\,dy = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dxdy$$

Converting to polar coordinates with $x = r\cos\theta, y = r\sin\theta$:

$$I^2 = \iint_{\mathbb{R}^2} e^{-r^2}r\,drd\theta = \int_0^{2\pi} d\theta \int_0^{\infty} re^{-r^2}\,dr$$

For the inner integral, use substitution $u = r^2, du = 2rdr$:

$$\int_0^{\infty} re^{-r^2}\,dr = \frac{1}{2}\int_0^{\infty} e^{-u}\,du = \frac{1}{2}[-e^{-u}]_0^{\infty} = \frac{1}{2}$$

Therefore:

$$I^2 = \int_0^{2\pi} d\theta \cdot \frac{1}{2} = \pi$$

Taking the square root of both sides: $I = \sqrt{\pi}$

Hence, $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$

Now, to show that $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$ is a density function:

Clearly $f(x) > 0$ for all $x$ when $\sigma > 0$

We need to verify that $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx$$

Using substitution $z = \dfrac{x-\mu}{\sigma}, dx = \sigma dz$:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{z^2}{2}\right\}\sigma dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}dz$$

Further substitution with $w = \dfrac{z}{\sqrt{2}}, dz = \sqrt{2}dw$:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-w^2}\sqrt{2}dw = \frac{\sqrt{2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-w^2}dw = \frac{\sqrt{2}}{\sqrt{2\pi}}\cdot\sqrt{\pi} = 1$$

Therefore, $f(x)$ is indeed a valid probability density function when $\sigma > 0$.

(b)

For a standard normal random variable $Z$ with PDF $f(z) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$, we calculate:

Mean calculation:

$$E[Z] = \int_{-\infty}^{\infty} z \cdot f(z)\,dz = \int_{-\infty}^{\infty} z \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\,dz$$

$$E[Z] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z \cdot e^{-\frac{z^2}{2}}\,dz$$

Note that $z \cdot e^{-\frac{z^2}{2}}$ is odd since $z$ is odd and $e^{-\frac{z^2}{2}}$ is even

Therefore, the integral over the symmetric interval $(-\infty, \infty)$ equals zero

$$E[Z] = 0$$

Variance calculation:

$$\text{Var}(Z) = E[Z^2] - (E[Z])^2 = E[Z^2] - 0^2 = E[Z^2]$$

$$E[Z^2] = \int_{-\infty}^{\infty} z^2 \cdot f(z)\,dz = \int_{-\infty}^{\infty} z^2 \cdot \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\,dz$$

$$E[Z^2] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2 \cdot e^{-\frac{z^2}{2}}\,dz$$

Using integration by parts with $u = z$ and $dv = z \cdot e^{-\frac{z^2}{2}}\,dz$:

We get $du = dz$ and $v = -e^{-\frac{z^2}{2}}$, thus:

$$E[Z^2] = \frac{1}{\sqrt{2\pi}}\left[-z \cdot e^{-\frac{z^2}{2}}\right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\,dz$$

The first term equals 0, and the second term equals 1 (integral of standard normal PDF)

$E[Z^2] = 1$

Therefore: $\text{Var}(Z) = E[Z^2] - (E[Z])^2 = 1 - 0 = 1$

(c)

Let's define $f(y) = y^{-1}e^{-\frac{1}{2}y^2}$ and $g(y) = (y^{-1} - y^{-3})e^{-\frac{1}{2}y^2}$

Computing the derivatives:

$f'(y) = -(1 + y^{-2})e^{-\frac{1}{2}y^2}$ and $g'(y) = -(1 - 3y^{-4})e^{-\frac{1}{2}y^2}$

Since $1 - 3y^{-4} < 1 < 1 + y^{-2}$ for $y > 0$

We have $-f'(y) > -g'(y)$, thus $f'(y) < g'(y)$

Multiplying by $\dfrac{1}{\sqrt{2\pi}}$ and integrating from $x$ to $\infty$:

$$\int_x^\infty \frac{1}{\sqrt{2\pi}}f'(y)dy < \int_x^\infty \frac{1}{\sqrt{2\pi}}g'(y)dy$$

Since $f(\infty) = g(\infty) = 0$, we get $\dfrac{1}{\sqrt{2\pi}}f(x) > \dfrac{1}{\sqrt{2\pi}}g(x)$

For the upper bound, note that:

$$\sqrt{2\pi}[1 - \Phi(x)] = \int_x^\infty e^{-\frac{1}{2}t^2}dt < \int_x^\infty \frac{x}{t}e^{-\frac{1}{2}t^2}dt$$

$$= x\int_x^\infty \frac{1}{t}e^{-\frac{1}{2}t^2}dt = x^{-1}e^{-\frac{1}{2}x^2}$$

For the lower bound, using integration by parts:

$$\sqrt{2\pi}[1 - \Phi(x)] = \int_x^\infty e^{-\frac{1}{2}t^2}dt$$

$$= \frac{e^{-\frac{1}{2}x^2}}{x} + \int_x^\infty \frac{e^{-\frac{1}{2}t^2}}{t^2}dt > \frac{e^{-\frac{1}{2}x^2}}{x} - \frac{e^{-\frac{1}{2}x^2}}{x^3}$$

$$= (x^{-1} - x^{-3})e^{-\frac{1}{2}x^2}$$

Therefore: $(x^{-1} - x^{-3})e^{-\frac{1}{2}x^2} < \sqrt{2\pi}[1 - \Phi(x)] < x^{-1}e^{-\frac{1}{2}x^2}, \quad x > 0$

(d)

Define $\alpha(x) = \mathbb{P}(X > x + a/x \mid X > x) = \dfrac{1 - \Phi(x + a/x)}{1 - \Phi(x)}$

From part (c), we know that for large $t > 0$:

$(t^{-1} - t^{-3})e^{-\frac{1}{2}t^2} < \sqrt{2\pi}[1 - \Phi(t)] < t^{-1}e^{-\frac{1}{2}t^2}$

As $x \to \infty$, we can approximate:

$1 - \Phi(x) \approx \dfrac{1}{\sqrt{2\pi}}x^{-1}e^{-\frac{1}{2}x^2}(1 + o(1))$

$1 - \Phi(x + a/x) \approx \dfrac{1}{\sqrt{2\pi}}(x + a/x)^{-1}e^{-\frac{1}{2}(x+a/x)^2}(1 + o(1))$

Therefore:

$$\alpha(x) \approx (1 + o(1)) \frac{x}{x + a/x} \cdot \frac{e^{-\frac{1}{2}(x+a/x)^2}}{e^{-\frac{1}{2}x^2}}$$

As $x \to \infty$, $\dfrac{x}{x + a/x} \to 1$

For the exponent term:

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

As $x \to \infty$, $-\dfrac{1}{2}(2a + a^2/x^2) \to -a$

Thus:

$$\alpha(x) = (1 + o(1)) \frac{e^{-\frac{1}{2}(x+a/x)^2}}{e^{-\frac{1}{2}x^2}} \to e^{-a} \text{ as } x \to \infty$$

$\square$

## Problem 6

(a) Let $X$ have a continuous distribution function $F$. Show that

(i) $F(X)$ is uniformly distributed on $[0, 1]$,

(ii) $-\log F(X)$ is exponentially distributed.

(b) A straight line $l$ touches a circle with unit diameter at the point P which is diametrically opposed on the circle to another point Q. A straight line QR joins Q to some point R on $l$. If the angle $\widehat{PQR}$ between the lines PQ and QR is a random variable with the uniform distribution on $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$, show that the length of PR has the Cauchy distribution (this length is measured positive or negative depending upon which side of P the point R lies).

(c) Let the net scores in the two halves of a game between teams $A$ and $B$ be independent and identically distributed random variables $X, Y$ that are symmetric about 0 and have a continuous distribution function $F$. Team $A$ wins (respectively, loses) if $X + Y > 0$ (respectively, $X + Y \le 0$). Find the probability that $A$ wins conditional on the half-time score $X$.

*Proof.*

(i) $F(X)$ is uniformly distributed on $[0, 1]$,

For any $0 \le y \le 1$, we need to find $\mathbb{P}(F(X) \le y)$ :

$\mathbb{P}(F(X) \le y) = \mathbb{P}(X \le F^{-1}(y))$

Since $F$ is the distribution function of $X$, we have:

$\mathbb{P}(X \le F^{-1}(y)) = F(F^{-1}(y)) = y$

This holds for all $y \in [0, 1]$, so $F(X) \sim \text{Uniform}[0, 1]$

(ii) $-\log F(X)$ is exponentially distributed.

Let $Y = -\log F(X)$. We need to show $Y \sim \text{Exp}(1)$

For any $t \geq 0$, we have:

$\mathbb{P}(Y \leq t) = \mathbb{P}(-\log F(X) \leq t) = \mathbb{P}(F(X) \geq e^{-t})$

Since $F(X) \sim \text{Uniform}[0, 1]$ from part (i), we have:

$\mathbb{P}(F(X) \geq e^{-t}) = 1 - \mathbb{P}(F(X) < e^{-t}) = 1 - e^{-t}$

Therefore, $\mathbb{P}(Y \leq t) = 1 - e^{-t}$ for $t \geq 0$

This is precisely the CDF of an exponential distribution with rate parameter 1

Thus, $-\log F(X) \sim \text{Exp}(1)$

(b)

Let's place the center of the circle at the origin, with P at $(0, \frac{1}{2})$ and Q at $(0, -\frac{1}{2})$.

The line $l$ is horizontal at $y = \frac{1}{2}$ (tangent to the circle at P).

Let $\theta = \widehat{PQR}$ with $\theta \sim \text{Uniform}[-\frac{\pi}{2}, \frac{\pi}{2}]$.

The line QR has slope $-\dfrac{1}{\tan(\theta)}$ with equation:

$y = -\dfrac{1}{2} - \dfrac{x}{\tan(\theta)}$

Intersecting with $l$ where $y = \dfrac{1}{2}$:

$\dfrac{1}{2} = -\dfrac{1}{2} - \dfrac{x}{\tan(\theta)} \Rightarrow x = -\tan(\theta)$

So R has coordinates $(-\tan(\theta), \frac{1}{2})$.

Considering the sign, $PR = -\tan(\theta)$ (negative when R is left of P, positive when right)

For any $t$, the CDF is:

$\mathbb{P}(PR \leq t) = \mathbb{P}(-\tan(\theta) \leq t) = \mathbb{P}(\theta \geq \arctan(-t))$

$= \dfrac{\frac{\pi}{2} - \arctan(-t)}{\pi} = \dfrac{1}{2} + \dfrac{\arctan(t)}{\pi}$

This is the CDF of the standard Cauchy distribution, thus PR has the Cauchy distribution.

(c)

Let's find the probability that team A wins given the half-time score $X = x$.

$\mathbb{P}(X + Y > 0 \mid X = x) = \mathbb{P}(Y > -x \mid X = x)$

Since $X$ and $Y$ are independent random variables:

$\mathbb{P}(Y > -x \mid X = x) = \mathbb{P}(Y > -x)$

Now, using the symmetry of $Y$ about 0:

$\mathbb{P}(Y > -x) = \mathbb{P}(Y < x)$    (by symmetry of the distribution)

$\mathbb{P}(Y < x) = F(x)$    (by definition of the CDF)

Therefore:

$\mathbb{P}(X + Y > 0 \mid X = x) = F(x)$

Furthermore, note that $F(X)$ is uniformly distributed on (0,1) by part (a)(i),

so the conditional probability of team A winning has mean $\dfrac{1}{2}$.

$\square$

## Problem 7

Secretary/marriage problem. You are permitted to inspect the $n$ prizes at a fête in a given order, at each stage either rejecting or accepting the prize under consideration. There is no recall, in the sense that no rejected prize may be accepted later. It may be assumed that, given complete information, the prizes may be ranked in a strict order of preference, and that the order of presentation is independent of this ranking. Find the strategy which maximizes the probability of accepting the best prize, and describe its behaviour when $n$ is large.

*Proof.* Since the order is random, the probability that the best candidate is in position $i$ is

$$\frac{1}{n}, \quad i = 1, 2, \ldots, n.$$

If the best candidate appears in position $i$, where $i > r$, we will choose it if and only if the best candidate among the first $i-1$ candidates appears in the first $r$ positions. Since every permutation of the first $i-1$ candidates is equally likely, the probability that the highest-ranked candidate among the first $i-1$ candidates is among the first $r$ is

$$\frac{r}{i-1}.$$

Therefore, the overall probability of selecting the best candidate is

$$P_n(r) = \sum_{i=r+1}^{n} \Pr\Big(\text{Best candidate is in position } i\Big) \cdot \Pr\Big(\text{Best among the first } i-1 \text{ is in the first } r\Big)$$

$$= \sum_{i=r+1}^{n} \frac{1}{n} \cdot \frac{r}{i-1}$$

$$= \frac{r}{n} \sum_{i=r+1}^{n} \frac{1}{i-1}.$$

Letting $j = i - 1$, the summation becomes

$$\sum_{j=r}^{n-1} \frac{1}{j}.$$

Thus,

$$P_n(r) = \frac{r}{n} \sum_{j=r}^{n-1} \frac{1}{j}.$$

For large values of $n$, the sum can be approximated by an integral:

$$\sum_{j=r}^{n-1} \frac{1}{j} \approx \int_r^n \frac{1}{x} \, dx$$

$$= \ln(n) - \ln(r)$$

$$= \ln\left(\frac{n}{r}\right).$$

Hence, for large $n$,

$$P_n(r) \approx \frac{r}{n} \ln\left(\frac{n}{r}\right).$$

To maximize $P_n(r)$ with respect to the cutoff $r$, consider $r$ as a continuous variable and define

$$P(r) = \frac{r}{n} \ln\left(\frac{n}{r}\right).$$

Differentiate $P(r)$ with respect to $r$:

$$\frac{dP}{dr} = \frac{1}{n} \ln\left(\frac{n}{r}\right) + \frac{r}{n} \left(-\frac{1}{r}\right)$$

$$= \frac{1}{n} \ln\left(\frac{n}{r}\right) - \frac{1}{n}.$$

Setting the derivative equal to zero to find the maximum, we obtain:

$$\frac{1}{n} \ln\left(\frac{n}{r}\right) - \frac{1}{n} = 0,$$

$$\ln\left(\frac{n}{r}\right) = 1,$$

$$\frac{n}{r} = e,$$

$$r = \frac{n}{e}.$$

Substitute $r = \frac{n}{e}$ back into the approximate expression for $P_n(r)$:

$$P_n\left(\frac{n}{e}\right) \approx \frac{\frac{n}{e}}{n} \ln\left(\frac{n}{n/e}\right)$$

$$= \frac{1}{e} \ln(e)$$

$$= \frac{1}{e}.$$

Thus, the optimal strategy is to reject the first $\frac{n}{e}$ candidates and then select the first candidate who is better than every candidate seen in the observation phase. For large $n$, this strategy yields a probability of approximately $\frac{1}{e} \approx 36.8\%$ of selecting the best candidate.

$\square$