HW3

1. **On the multivariate normal distribution** [4 Points]

Let $X = (X_1, X_2)$ be a random vector with $X \sim \mathcal{N}_2(\mu, \Sigma)$, where

$$\mu = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}.$$

(a) Write down the probability density function of the law of X.

(a)

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

Here. $\Sigma^{-1} = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$ $\quad \det(\Sigma) = 5$

$$\Rightarrow f_X\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2\sqrt{5}\,\pi} \exp\left(-\frac{1}{2}\left(\begin{pmatrix} x_1-1 \\ x_2+2 \end{pmatrix}^T \cdot \frac{1}{5}\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_1-1 \\ x_2+2 \end{bmatrix}\right)\right)$$

$$= \frac{1}{2\sqrt{5}\pi} \exp\left(-\frac{1}{10}\left(2x_1^2 - 2x_1x_2 + 3x_2^2 - 8x_1 + 14x_2 + 18\right)\right)$$

(b) Find the distribution of $Y = X_1 - 3X_2$ and of $Z = (X_1 + X_2, X_1 - X_2)$.

(b)

$$Y = \underbrace{[1 \ -3]}_{:=A} X$$

$$\Rightarrow Y \sim \mathcal{N}(A\mu, A\Sigma A^T) = \mathcal{N}(7, 15)$$

$$Z = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{:=B} X$$

$$\Rightarrow Z \sim \mathcal{N}(B\mu, B\Sigma B^T) = \mathcal{N}\left(\begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 7 & 1 \\ 1 & 3 \end{bmatrix}\right)$$

(c) Consider $U \sim \mathcal{N}(0,1)$ and $R \sim Ber(\frac{1}{2})$, independent from U. Set

$$V = \begin{cases} U, & R = 1, \\ -U, & R = 0. \end{cases}$$

It can be proved that $V \sim \mathcal{N}(0,1)$.[1] Show that $\mathrm{Cov}[U,V] = 0$, and that U and V are *not* independent. Why is this not a contradiction to Remark 2.3 (i)?

$V \sim N(0,1)$

proof of $V \sim N(0,1)$

$$\begin{aligned} F_V(x) = \mathbb{P}(V \le x) &= \mathbb{P}(V \le x \mid R=1)\cdot \mathbb{P}(R=1) + \mathbb{P}(V \le x \mid R=0)\,\mathbb{P}(R=0) \\ &= \tfrac{1}{2}\cdot \mathbb{P}(U \le x) + \tfrac{1}{2}\,\mathbb{P}(-U \le x) \\ &= \tfrac{1}{2}\cdot \mathbb{P}(U \le x) + \tfrac{1}{2}\,\mathbb{P}(U \ge -x) \\ &= \tfrac{1}{2}\cdot \mathbb{P}(U \le x) + \tfrac{1}{2}\,\mathbb{P}(U \le x) \\ &= \mathbb{P}(U \le x) \\ &= \Phi(x) \end{aligned}$$

Here we show $F_V(x)$ is the distribution of Φ.

i.e. $V \sim N(0,1)$

$\mathrm{Cov}(U,V)$

$$\mathrm{Cov}(U,V) = \mathbb{E}[UV] - \mathbb{E}[U]\cdot\mathbb{E}[V]$$

$$\mathbb{E}[UV] = \int x^2 f(x)\cdot \tfrac{1}{2}\,dx + \int -x^2 f(x)\cdot \tfrac{1}{2}\,dx = 0$$

$$\Rightarrow \mathrm{Cov}(U,V) = \mathbb{E}[UV] - \mathbb{E}[U]\cdot\mathbb{E}[V] = 0 - 0\cdot 0 = 0$$

Dependence

Let $a<0$

$$\mathbb{P}(U \le a)\cdot\mathbb{P}(V \le a) = \left(\int_{-\infty}^{a} f(x)\,dx\right)^2 \quad \text{where } f \text{ is the density of } N(0,1)$$

$$\mathbb{P}(U \le a \cap V \le a) = \tfrac{1}{2}\int_{-\infty}^{a} f(x)\,dx + \tfrac{1}{2}\left(\int_{-\infty}^{a} f(x)\,dx\right)^2$$

Clearly: $\mathbb{P}(U \le a)\cdot\mathbb{P}(V \le a) \ne \mathbb{P}(U \le a \cap V \le a)$
which shows dependency.

Not a Contradiction.

U and V are not jointly distributed as multivariate normal.

2. **On the χ^2-distribution** [4 Points]

(a) Calculate the density of a χ^2-distribution with one degree of freedom. For this, recall if $X \sim \mathcal{N}(0,1)$, then $Z = X^2 \sim \chi_1^2$.
Hint: Look at $F_Z(z) = \mathbf{P}[X^2 \le z]$.

$z \geq 0$

(a)

$$F_Z(z) = \mathbb{P}[X^2 \le z]$$
$$= \mathbb{P}[-\sqrt{z} \le X \le \sqrt{z}] = \int_{-\sqrt{z}}^{\sqrt{z}} f(x)\, dx = \Phi(\sqrt{z}) - \Phi(-\sqrt{z})$$

$$f_Z(z) = F_Z'(z) = \frac{1}{2\sqrt{z}} f(\sqrt{z}) + \frac{1}{2\sqrt{z}} f(-\sqrt{z})$$
$$= \frac{1}{\sqrt{z}} f(\sqrt{z}) = \frac{1}{\sqrt{z}} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\sqrt{z})^2}{2}\right)$$

Therefore. $f_Z(z) = \dfrac{1}{\sqrt{2\pi\, z}} \cdot \exp\left(-\dfrac{z}{2}\right) \cdot \mathbb{1}_{(0,+\infty)}$

(b)

(b) Verify that the χ^2-distribution with n degrees of freedom has density

$$f_{\chi_n^2}(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} \mathbb{1}_{(0,\infty)}(x).$$

Hint: You may use without proof the fact that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Then identify the density of χ_1^2 with a Γ-distribution and use the result from Problem 2 (b) on Problem Set 1.

$$\Rightarrow f_{\chi_1^2}(x) = \frac{1}{\sqrt{2}\ \Gamma(\frac{1}{2})} e^{-\frac{x}{2}} x^{-\frac{1}{2}} \cdot \mathbb{1}_{(0,+\infty)}(x)$$

Here we could see. $\chi_1^2 \sim \Gamma(\alpha, \beta)$ where $\alpha = \frac{1}{2}$ $\beta = \frac{1}{2}$
$\sim \Gamma(\frac{1}{2}, \frac{1}{2})$

Therefore. $\chi_n^2 = \sum_{i=1}^{n} X_i^2 \sim \sum_{i=1}^{n} \Gamma(\frac{1}{2}, \frac{1}{2})$

Therefore. by problem 2b in HW1

$$\chi_n^2 = \sum_{i=1}^{n} X_i^2 \sim \Gamma(\tfrac{n}{2}, \tfrac{1}{2})$$

Thus. $$f_{\chi_n^2} = \frac{x^{\frac{n}{2}-1} e^{-\frac{1}{2}x} (\frac{1}{2})^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \cdot \mathbb{1}_{(0,+\infty)}$$

$$= \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{1}{2}x} x^{\frac{n}{2}-1} \cdot \mathbb{1}_{(0,+\infty)}$$

(c) Suppose the coordinates (X_1, X_2, X_3) of a particle undergoing diffusive motion can be described at some time t by i.i.d. $\mathcal{N}(0,1)$-distributed random variables. What is the probability that the particle at time t is located within a ball of radius $\frac{1}{2}$?
Hint: Use freely available online calculators for the cumulative distribution function of the χ^2 distribution.

(C) At time t is located within a ball of radius $\frac{1}{2}$.

$\Leftrightarrow$

$\mathbb{P}(X_1^2 + X_2^2 + X_3^2 \le \frac{1}{4})$

$= \mathbb{P}(\chi_3^2 \le \frac{1}{4}) = F_{\chi_3^2}(\frac{1}{4}) = 0.03086$

3. **Derivation of the density of a t-distribution** [4 Points]

In this problem, we show step-by-step that the density of the t-distribution with n degrees of freedom is given by

$$f_{t_n}(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

(a) Recall that $T = \frac{X}{\sqrt{Y/n}}$ has the t_n-distributiuon if $X \sim \mathcal{N}(0,1)$ and $Y \sim \chi_n^2$ are independent. Write down the probability density function of $\sqrt{Y/n}$.
Hint: Look at $F_{\sqrt{Y/n}}(y) = \mathbf{P}[\sqrt{Y/n} \le y]$. The density of Y is known from Problem 2 (b).

(a) if $y > 0$

$$F_{\sqrt{Y/n}}(y) = \mathbb{P}[\sqrt{Y/n} \le y] = \mathbb{P}[Y \le ny^2] = F_{\chi_n^2}(ny^2)$$

$$f_{\sqrt{Y/n}} = F'_{\sqrt{Y/n}}(y) = 2ny\, F'_{\chi_n^2}(ny^2) = 2ny \cdot \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{1}{2}ny^2} (ny^2)^{\frac{n}{2}-1}$$

$$\Rightarrow f_{\sqrt{Y/n}} = 2ny \cdot \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\frac{1}{2}ny^2} (ny^2)^{\frac{n}{2}-1} \cdot \mathbb{1}_{[0,+\infty)}$$

(b) For the quotient of two independent, continuous, real random variables X and Y where Y is positive, one can also show that $Z = \frac{X}{Y}$ has density

$$f_Z(z) = \int_0^\infty y f_X(zy) f_Y(y)\mathrm{d}y,\ z \in \mathbb{R}.$$

Use this in combination with (a) to demonstrate that

$$(\star) \qquad f_{t_n}(z) = \frac{2^{1-\frac{n}{2}}}{\sqrt{2\pi}\,\Gamma(\frac{n}{2})} n^{\frac{n}{2}} \int_0^\infty y^n e^{-\frac{1}{2}(n+z^2)y^2}\mathrm{d}y.$$

proof. $$f_{t_n(z)}(z) = f_{\frac{X}{\sqrt{Y/n}}} = \int_0^{+\infty} y\, e^{-\frac{1}{2}z^2y^2}\frac{1}{\sqrt{2\pi}} \cdot 2ny \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \cdot e^{-\frac{1}{2}ny^2} (ny^2)^{\frac{n}{2}-1} dy$$

$$= \frac{2^{1-\frac{n}{2}}}{\sqrt{2\pi}\,\Gamma(\frac{n}{2})} n^{\frac{n}{2}} \int_0^{+\infty} y^n e^{-\frac{1}{2}(n+z^2)y^2} dy$$

(c) Recall now that the Γ-function is defined for $\alpha > 0$ by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{d}x.$$

Validate the formula for f_{t_n} by using a substitution in the remaining integral in (b).

As $$f_{t^n}(z) = \frac{2^{1-\frac{n}{2}}}{\sqrt{2\pi}\,\Gamma(\frac{n}{2})} n^{\frac{n}{2}} \int_0^{+\infty} y^n e^{-\frac{1}{2}(n+z^2)y^2} dy$$

$$\int_0^{+\infty} y^n e^{-\frac{1}{2}(n+z^2)y^2} dy$$

let $u(y) = \frac{1}{2}(n+z^2)y^2 \quad u'(y) = (n+z^2)y$

let $\alpha = \frac{n+1}{2}$

$$\Rightarrow \int_0^{+\infty} y^n e^{-\frac{1}{2}(n+z^2)y^2}$$

$$= \left[\tfrac{1}{2}(n+z^2)\right]^{-\frac{n+1}{2}} \int_0^{+\infty} u(y)^{\frac{n+1}{2}-1} e^{-u(y)} \cdot u'(y)\, dy$$

$$= \frac{1}{2}\left[\tfrac{1}{2}(n+z^2)\right]^{-\frac{n+1}{2}} \int_0^{+\infty} u^{\frac{n+1}{2}-1} e^{-u} du$$

$$= \frac{1}{2}\left[\tfrac{1}{2}(n+z^2)\right]^{-\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right)$$

$$\Rightarrow f_{t^2}(z) = \frac{2^{1-\frac{n}{2}}}{\sqrt{2\pi}\,\Gamma(\frac{n}{2})} n^{\frac{n}{2}} \cdot \frac{1}{2}\left[\tfrac{1}{2}(n+z^2)\right]^{-\frac{n+1}{2}} \cdot \Gamma\left(\frac{n+1}{2}\right)$$

$$= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})} \left(1 - \frac{z^2}{n}\right)^{-\frac{n+1}{2}}$$

4. **Properties of estimators** [4 Points]

Let $X_1, ..., X_n$ be i.i.d. real random variables with $X_1 \sim Pois(\theta)$, with $\theta > 0$. We want to estimate $\gamma = \mathbf{P}_\theta[X_1 = 0] = e^{-\theta}$ based on the data $X_1, ..., X_n$. We consider the two estimators for γ:

$$\widehat{\gamma}_1 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i=0\}}, \qquad \widehat{\gamma}_2 = \exp\left(-\frac{1}{n}\sum_{i=1}^{n} X_i\right).$$

(a) Show that $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ are both consistent estimators for γ.
Hint: Use the weak law of large numbers and the continuous mapping theorem.

Proof. To show they're consistent,

$$\hat{\gamma}_1(X_1, \cdots, X_n) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{X_i = 0\}$$

$$\hat{\gamma}_1 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{X_i = 0\}$$

By Law of large numbers. $\hat{\gamma}_1 \xrightarrow{\mathbb{P}_\theta} \mathbb{E}[\mathbb{1}\{X_i = 0\}] = \mathbb{P}_\theta[X_1 = 0] = e^{-\theta}$

which shows $\hat{\gamma}_1$ is a consistent estimator for γ.

$$\hat{\gamma}_2 = \exp\left(-\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

By law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}_\theta} \theta$$

By Continuous Mapping Theorem,

$$\hat{\gamma}_2 = \exp\left(-\frac{1}{n}\sum_{i=1}^{n} X_i\right) \xrightarrow{\mathbb{P}_\theta} \exp(-\theta) = \gamma$$

which shows $\hat{\gamma}_2$ is a consistent estimator.

(b) Determine whether $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ are unbiased.

(b)

$$\mathbb{E}(\hat{\gamma}_1) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{x_i=0\}\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\mathbb{1}\{x_i=0\}\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{P}_\theta(X_1=0)$$

$$= e^{-\theta} = e^{-\theta}$$

$$\mathbb{E}(\hat{\gamma}_2) = \mathbb{E}\left[\exp\left(-\frac{1}{n}\sum x_i\right)\right] = \mathbb{E}\left[\left(\prod_{i=1}^{n} \exp\left(-\frac{1}{n}x_i\right)\right)\right]$$

$$\overset{iid}{=} \prod_{i=1}^{n}\mathbb{E}\left[\exp\left(-\frac{1}{n}x_i\right)\right]$$

$$= \left(\mathbb{E}\left[\exp\left(-\frac{1}{n}x_i\right)\right]\right)^n = \left(\sum_{k=0}^{\infty}\exp\left(-\frac{1}{n}k\right)\frac{\theta^k e^{-\theta}}{k!}\right)^n$$

$$= \left(e^{-\theta}\sum_{k=0}^{\infty}\frac{[\theta\cdot\exp(-\frac{1}{n})]^k}{k!}\right)$$

$$= e^{-\theta}\cdot e^{\theta\cdot\exp(-\frac{1}{n})}$$

$$= e^{\theta(\exp(-\frac{1}{n})-1)} \neq e^{-\theta}$$

$\hat{\gamma}_1$ Unbiased $\hat{\gamma}_2$ biased.

(c) Calculate the mean square error $\text{MSE}_\theta(\widehat{\gamma}_1) = \mathbf{E}_\theta\left[(\widehat{\gamma}_1 - \gamma)^2\right]$.

(C).

$$\text{MSE}(\hat{\gamma}_1) = \mathbb{E}_\theta\left[(\hat{\gamma}_1 - \gamma)^2\right]$$

As $\hat{\gamma}_1$ is unbiased

$$\text{MSE}[\hat{\gamma}_1] = \text{Var}[\hat{\gamma}_1] = \text{Var}\left[\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\{x_i=0\}\right]$$

$$= \frac{1}{n^2}\cdot n\cdot\text{Var}\left[\mathbb{1}\{x_i=0\}\right]$$

$$= \frac{1}{n}\left(\mathbb{E}\left[\mathbb{1}^2\{x_i=0\}\right] - \mathbb{E}^2\left[\mathbb{1}\{x_i=0\}\right]\right)$$

$$= \frac{1}{n}\left(e^{-\theta} - e^{-2\theta}\right) = \frac{1}{n}e^{-\theta}(1-e^{-\theta})$$