Solution.

CONCORDIA UNIVERSITY
Department of Economics

ECON 222 SECTIONS A, B
STATISTICAL METHODS II
Fall 2017 – MIDTERM 1
Sunday, October 15, 11:00am – 13:00pm

1. **(8 marks)** Find df/dx for the following functions.

a. **(2 marks)** $f(x) = \ln(x^2 + x)$

$$\frac{df}{dx} = \frac{d[x^2+x]/dx}{x^2+x} = \frac{2x+1}{x^2+x}$$

b. **(2 marks)** $f(x) = e^{x^2}$

$$\frac{df}{dx} = 2x\,e^{x^2}$$

c. **(2 marks)** $f(x) = x^2 \ln(x)$

$$\frac{df(x)}{dx} = 2x\ln(x) + x^2\frac{1}{x}$$
$$= 2x\ln(x) + x = x(2\ln(x) + 1)$$

d. **(2 marks)** $f(x) = e^x/x$

$$\frac{df(x)}{dx} = \frac{xe^x - e^x}{x^2} = \frac{e^x(x-1)}{x^2}$$

2. **(6 marks)** The random variable $W = aX + bY + cZ$ is a linear combination of three other random variables, X, Y and Z, where $E(X) = \mu_X$, $\text{var}(X) = \sigma_X^2$, $E(Y) = \mu_Y$, $\text{var}(Y) = \sigma_Y^2$, $E(Z) = \mu_Z$, and $\text{var}(Z) = \sigma_Z^2$.

a. **(3 marks)** Find $E(W)$

$$E(w) = E(ax + by + cz)$$
$$= aE(x) + bE(y) + cE(z)$$
$$= a\mu_x + b\mu_y + c\mu_z$$

b. **(3 marks)** Find ~~var(Z)~~ $Var(W) = Var(ax + by + cz)$

$$= a^2 Var(x) + b^2 Var(y) + c^2 Var(z) + 2ab\, Cov(x,y) + 2ac\, Cov(x,z) + 2bc\, Cov(y,z)$$

$$= a^2\sigma_x + b^2\sigma_y + c^2\sigma_z + 2ab\sigma_{xy} + 2ac\sigma_{xz} + 2bc\,\sigma_{yz}.$$

3. **(18 marks)** Let X and Y be continuous random variables with a joint probability density function (PDF) of $f(x,y) = 4xy$ for $x \in [0,1]$ and $y \in [0,1]$. Given that $f(x,y) \geq 0$ for all $x \in [0,1]$ and $y \in [0,1]$:

a. **(3 marks)** Show that $f(x,y)$ is a probability density function.

$$\int_0^1\int_0^1 4xy\,dx\,dy = \int_0^1 4y\left[\int_0^1 x\,dx\right]dy = \int_0^1 4y\left[\frac{x^2}{2}\right]_0^1 dy$$

$$= \int_0^1 4y\left(\frac{1}{2}\right)dy = \int_0^1 2y\,dy = y^2\Big|_0^1 = 1$$

$\Longrightarrow f(x,y)$ is a probability density function

b. (3 marks) Calculate $\Pr\left(0 \le X \le \frac{1}{2}, \frac{1}{4} \le Y \le \frac{3}{4}\right)$.

$$\int_{1/4}^{3/4}\int_{0}^{1/2} 4xy\,dx\,dy = \int_{1/4}^{3/4} 4y\left[\frac{x^2}{2}\right]_0^{1/2} dy = \int_{1/4}^{3/4} 4y\left[\frac{1}{8}\right] dy = \int_{1/4}^{3/4} \frac{1}{2} y\,dy$$

$$= \frac{1}{2}\left[\frac{y^2}{2}\right]_{1/4}^{3/4} = \frac{1}{2}\left[\frac{3.3}{4.4.2} - \frac{1}{4.4.2}\right] = \frac{1}{2}\left\{\frac{8}{32}\right\} = \frac{4}{32}$$

c. (3 marks) Find the marginal PDF of X.

$$\int_0^1 4xy\,dy = 4x\int_0^1 y\,dy = 4x\left[\frac{y^2}{2}\right]_0^1 = \frac{4x}{2}$$

$$\Rightarrow \quad f_X(x) = 2x \qquad 0 \le x \le 1$$

d. (3 marks) Calculate the expected value of X.

$$E(x) = \int_0^1 x(2x)\,dx = 2\int_0^1 x^2\,dx = 2\left[\frac{x^3}{3}\right]_0^1$$

$$\Rightarrow \quad E(x) = \frac{2}{3}$$

e. **(3 marks)** Calculate the variance of X.

$$E(X^2) = \int_0^1 x^2(2x)\,dx = 2\int_0^1 x^3\,dx$$

$$= 2\left[\frac{x^4}{4}\right]_0^1 = 2\left[\frac{1}{4}\right] = \frac{1}{2}$$

$$\Rightarrow Var(X) = E(X^2) - [E(X)]^2$$

$$= \frac{1}{2} - \left(\frac{2}{3}\right)^2$$

$$= \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

f. **(3 marks)** Find the conditional PDF of Y given X.

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{4xy}{2x} = 2y \qquad 0 \le y \le 1$$

4. **(8 marks)** Consider a random variable $V = Z_1^2 + Z_2^2 + \ldots + Z_n^2 \sim \chi_n^2$, where $Z_i \sim iid\, N(0,1)$.

a. **(4 marks)** Prove that $E(V) = n$. (Hint: $\text{var}(Z_i) = E(Z_i^2) - E(Z_i)^2$.)

$Var(Z_i) = 1 \qquad E(Z_i) = 0 \qquad \Rightarrow E(Z_i^2) = Var(Z_i) + E(Z_i)^2$

$$= 1 + 0$$

$$\Rightarrow E(Z_i^2) = 1$$

$$E(V) = E(Z_1^2) + E(Z_2^2) + \cdots + E(Z_n)^2$$

$$= 1 + 1 + \cdots + 1$$

$$= n(1)$$

$$\Rightarrow E(V) = n.$$

b. **(4 marks)** Prove that $\operatorname{var}(V) = 2n$.

(Hint: $\operatorname{var}(V) = E(V^2) - E(V)^2$ and $E(V^2) = E\left(\left(Z_1^2 + \ldots + Z_n^2\right)^2\right)$.)

$Var(V) = Var\left(\sum_{i=1}^{n} (Z_i^2)\right)$

Let $n = 2 \Rightarrow E(V^2) = E\left[(Z_1^2 + Z_2^2)^2\right] = E(Z_1^2 Z_1^2 + 2Z_1^2 Z_2^2 + Z_2^2 Z_2^2)$

$= E(Z_1^2 Z_1^2) + 2E(Z_1^2 Z_2^2) + E(Z_2^2 Z_2^2)$

$= \underbrace{E(Z_1^4)}_{3} + 2E(Z_1^2 Z_2^2) + \underbrace{E(Z_2^4)}_{3}$

$\boxed{\text{for a standard normal the fourth moment is } E(Z_i^4) = 3}$

$\Rightarrow$ and $\because Z_1$ & Z_2 are independent

$\Rightarrow E(Z_1^2 Z_2^2) = E(Z_1^2)E(Z_2^2) = 1$

$\Rightarrow E(V^2) = 2E(Z_1^2)E(Z_2^2) + 3 + 3$

$= 8$

in general $E(V^2) = \sum_{i=1}^{n} E(Z_i^2 Z_i)^2 + 2\sum_{i=1}^{n} \sum_{j=i}^{n} E(Z_i^2)E(Z_j^2)$

$= n(3) + 2\left(\frac{n^2 - n}{2}\right)$

$= 3n + n^2 - n = n^2 + 2n$

$\Rightarrow Var(V) = n^2 + 2n - n^2$

$\Rightarrow Var(V) = 2n$

if $n = 2$; $Var(V) = 8 - (2.2) = 4$.

5. **(10 marks)** A researcher draws a random sample Y_1, Y_2 and Y_3 from a $N(\mu, \sigma^2)$ population. She is deciding between two estimators of μ: $\bar{Y} = \frac{1}{3}Y_1 + \frac{1}{3}Y_2 + \frac{1}{3}Y_3$ and $\tilde{Y} = \frac{1}{2}Y_1 + \frac{1}{3}Y_2 + \frac{1}{6}Y_3$.

a. **(4 marks)** Prove that both $\bar{Y}$ and $\tilde{Y}$ are linear unbiased estimators.

$E(\bar{Y}) = \frac{1}{3}E(Y_1) + \frac{1}{3}E(Y_2) + \frac{1}{3}E(Y_3)$

$= \frac{1}{3}(\mu + \mu + \mu)$

$= \mu$

$\because \bar{Y} = \frac{1}{3}\sum_{i=1}^{3} Y_i$

$\Rightarrow$ it is linear in Y

$\because E(\bar{Y}) = \mu$

$\Rightarrow \bar{Y}$ is unbiased.

let $w_1 = \frac{1}{2}$, $w_2 = \frac{1}{3}$, $w_3 = \frac{1}{6}$

$\Rightarrow \tilde{Y} = \sum_{i=1}^{3} w_i Y_i$

$\Rightarrow \tilde{Y}$ is linear in Y

$E(\tilde{Y}) = \frac{1}{2}E(Y_1) + \frac{1}{3}E(Y_2) + \frac{1}{6}E(Y_3)$

$= \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{6}\right)\mu$

$= \frac{3 + 2 + 1}{6}\mu$

$= \mu$

$\because E(\tilde{Y}) = \mu \Rightarrow \tilde{Y}$ is unbiased

b. **(2 marks)** Calculate the variance of $\bar{Y}$.

$$Var(\bar{Y}) = \frac{1}{9}[\sigma^2] + \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2$$

$$= \frac{3\sigma^2}{9} = \frac{\sigma^2}{3}$$

c. **(2 marks)** Calculate the variance of $\tilde{Y}$.

$$Var(\tilde{Y}) = \frac{1}{4}\sigma^2 + \frac{1}{9}\sigma^2 + \frac{1}{36}\sigma^2$$

$$= \frac{9+4+1}{36}\sigma^2 = \frac{14}{36}\sigma^2$$

d. **(2 marks)** ***Briefly*** explain which estimator is better.

Both estimators are unbiased. However, $Var(\bar{Y}) = \frac{\sigma^2}{3}$ is less than $Var(\tilde{Y}) = \frac{14\sigma^2}{36}$

$\Longrightarrow$ $Var(\bar{Y})$ has a smaller variance

$\Longrightarrow$ $\bar{Y}$ is a better estimator.

6. **(20 marks)** An econometrician believes that stock returns are normally distributed with mean μ and variance σ^2 and she proceeded by collecting 1000 random observations. Let X be the stock return and $\bar{X} = \sum_1^N Xi/N$. Also note that $\sum_1^N Xi = 200$ and $\sum_1^N (Xi - \bar{X})^2 = 400$.

a. **(2 marks)** Suggest an unbiased estimator of (μ) and compute an estimate.

A natural estimator of the population mean is the sample average $\bar{X} = \frac{\sum X_i}{N}$. Note that $E(\bar{X}) = \mu \Rightarrow \bar{X}$ is unbiased.

$\bar{X} = \frac{\sum X_i}{N}$, $\sum X_i = 200$ $N = 1000 \Rightarrow \bar{X} = \frac{200}{1000} = 0.2$

b. **(2 marks)** Estimate the population variance (σ^2).

$$\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{N-1} = \frac{400}{999} = 0.4$$

c. **(6 marks)** Test the following with 99% confidence, explain your choice of the constructed statistic and demonstrate your answer graphically.
H$_0$: μ = 0.16
H$_1$: μ < 0.16

Because we do not know $\sigma^2 \Rightarrow$ we should use t-statistics.

$$\text{t-statistic} = \frac{\bar{X} - 0.16}{se(\bar{X})} = \frac{0.2 - 0.16}{\sqrt{0.4}} = 0.06324 \sim t_{(999)}$$

$\alpha = 1 - 0.99 = 0.01$

$P(t \le -t_c) = 0.01$

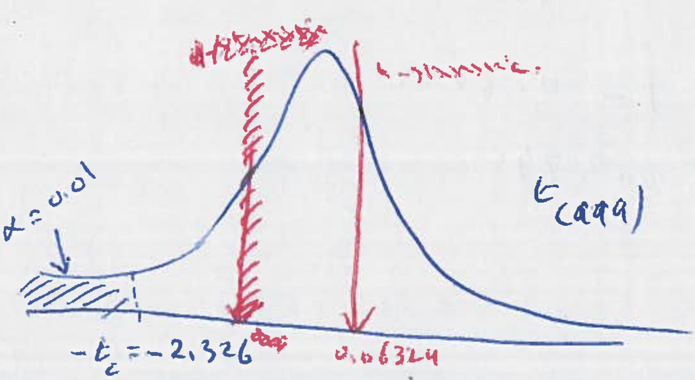
$\Rightarrow P(t \ge t_c) = 1 - P(t \le t_c) = 0.01$

$\Rightarrow P(t \le t_c) = 0.99$

$\Rightarrow t_c = 2.326$

$\because$ t-statistic $> (-t_c)$

$\Rightarrow$ we fail to reject H_0

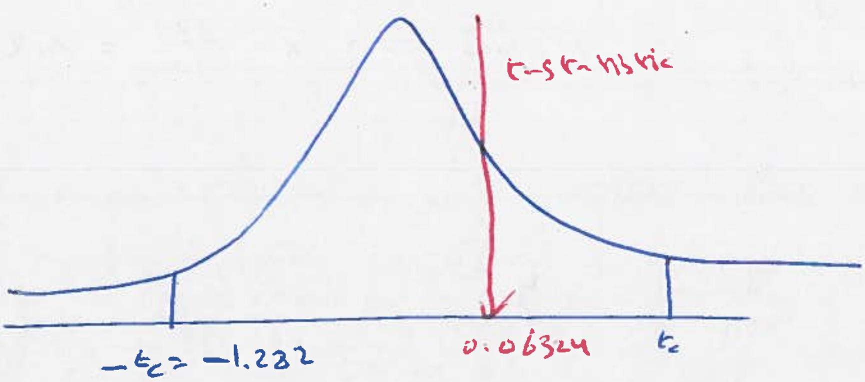


d. **(5 marks)** Would you reach to a different conclusion with 90% confidence level? Explain!

with 90% confidence, $\alpha = 1 - 0.9 = 0.10$

$\Rightarrow P(t \le -t_c) = P(t \ge t_c) = 1 - P(t \le t_c) = 0.10$

$\Rightarrow P(t \le t_c) = 0.9 \quad \Rightarrow t_c = 1.282$



$\therefore$ t-statistic $> (t_c)$

$\Rightarrow$ we fail to reject

$\Rightarrow$ we still would ~~reject~~ fail to reject with 90% confidence.

e. **(5 marks)** What is the probability of incurring a loss in the stock investment? [Hint: Losses occur when we have negative returns]

Losses would occur when $(X \le 0)$

$\Rightarrow P(X \le 0)$ is the answer.

Recall $X \sim N(\mu, \sigma^2)$

we could use $\bar{X}$ and $\hat{\sigma}^2 \Rightarrow X \sim N(0.2, 0.4)$

$P(X \le 0) = P\left(\frac{X - 0.2}{\sqrt{0.4}} \le \frac{0 - 0.2}{\sqrt{0.4}}\right) = P(Z \le -0.31623)$

$= 1 - P(Z \le 0.31623) = 1 - 0.6255$

$\simeq 0.3745$