CONCORDIA UNIVERSITY, DEPARTMENT OF ECONOMICS
ECON 222/4 - **STATISTICAL METHODS II** - WINTER 2017
MIDTERM 1 - Sunday, February 19 (11:00 – 13:00, H 937)

Student Name:

Student ID: Section:

Q1. **(12 marks)** Simplify the following expressions.

1A) **(8marks)** Find $\frac{dy}{dx}$, where

(i) $y = 3x^3 + 5x^2 + 9x$

$$\frac{dy}{dx} = 9x^2 + 10x + 9$$

(ii) $y = \ln(ax^2 + b)$

$$\frac{dy}{dx} = \frac{2ax}{ax^2+b}$$

(iii) $y = (x^2 + cx)(x + dx)$

$$\Rightarrow y = x^3 + dx^3 + cx^2 + cdx^2$$

$$\Rightarrow \frac{dy}{dx} = 3x^2 + 3dx^2 + 2cx + 2cdx \Rightarrow \frac{dy}{dx} = (3+3d)x^2 + (2c+2cd)x$$

(iv) $y = (7x^2 + 2x)/3x$ let $f(x) = 7x^2 + 2x$ and $g(x) = 3x \Rightarrow f'(x) = 14x + 2$, $g'(x) = 3$

$$\Rightarrow \frac{dy}{dx} = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2} = \frac{(14x+2)3x - (7x^2+2x)3}{9x^2} = \frac{30}{90}$$

1B) **(4 marks)** Simplify the following expressions.

(i) $\int_0^1 (4x^3)dx = x^4 \Big|_0^1 = 1$

(ii) $\int_0^1 2x(x^2+1)dx = 2\int_0^1 (x^3+x)\,dx = 2\left[\frac{x^4}{4} + \frac{x^2}{2}\right]_0^1 = 2\left[\frac{1}{4} + \frac{1}{2}\right] = \frac{3}{2}$

Q2. **(8 marks)** Consider X and Y are two random variables having a joint probability density function

2A) **(2 marks)** What does the ***Law of Iterated Expectations*** state about their relationship? Write its definition. Write the mathematical expression when X and Y are <u>discrete</u> random variables, and when X and Y are <u>continuous</u> random variables.

Law of iterated expectations states that the expected value of Y is the expected value of the conditional expectation of Y given X : $E(Y) = E[E(Y|X)]$

in the discrete case, this could be expressed as $E(Y) = \sum_x E(Y|x) f(x)$

and in the continuous case: $E(Y) = \int_X E(Y|X) f(x)\,dx$

2B) **(2 marks)** Prove the ***Law of Iterated Expectations*** for X and Y, assuming that X and Y are <u>continuous</u> random variables and have a joint probability density function ***(pdf)*** $f(x,y)$.

Lets start with the definition of $E(Y) = \int_y y f_y(y)\,dy$ where $f_y(y)$ is the marginal density of y.

$\Rightarrow E(Y) = \int_y y \left[\int_x f(x,y)\,dx\right] dy \quad \because f_y(y) = \int_x f(x,y)\,dx$

$= \int_y y \left[\int_x f(y|x) f_x(x)\,dx\right] dy \quad \because f(y|x) = \frac{f(x,y)}{f_x(x)} \Rightarrow f(x,y) = f(y|x) f_x(x)$

$= \int_x \left[\int_y y f(y|x)\,dy\right] f_x(x)\,dx$ [change the order of integration]

$\Rightarrow \boxed{E(Y) = \int_X E(Y|X) f_x(x)\,dx} \quad \because \int_y y f(y|x)\,dy = E(y|x)$

2C) **(4 marks)** Let X be a continuous random variable with a probability density function ***(pdf)*** $f(x) = 3x^3/8$ for $0 < x < 2$. Compute:

(i) $P(0 < X < ½)$ $= \frac{3}{8}\int_0^{1/2} x^3\,dx = \frac{3}{8}\left[\frac{x^4}{4}\right]_0^{1/2} = 0.005859$

(ii) $P(1 < X < 2)$ $= \frac{3}{8}\int_1^2 x^3\,dx = \frac{3}{8}\left[\frac{x^4}{4}\right]_1^2 = \frac{3}{8}\left[4 - \frac{1}{4}\right] = 1.5$

"This part is discarded, it will not be graded"

Q3. **(4 marks)** The random variable Z is a linear combination of two other random variables, X and Y,

$$Z = aX + bY \quad \text{where} \quad E(X) = \mu_X,\ \text{var}(X) = \sigma_X^2,\ E(Y) = \mu_Y \text{ and } \text{var}(Y) = \sigma_Y^2.$$

3A) **(2 marks)** Find the expectations of Z, *E(Z)*.

$$E(Z) = E(aX + bY)$$
$$= E(aX) + E(bY)$$
$$= a E(X) + b E(Y) \quad \because a, b \text{ are constants.}$$
$$\Rightarrow E(Z) = a\mu_x + b\mu_y.$$

3B) **(2 marks)** Find the variance of Z, Var(Z).

$$Var(Z) = Var(aX + bY)$$
$$= a^2 Var(x) + b^2 Var(Y) + 2ab\, Cov(X, Y)$$
$$= a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}^2 \quad \text{Let } \sigma_{xy}^2 = Cov(X, Y)$$

Q4. **(12 marks)** Suppose X and Y are continuous random variables with a joint probability density function (*pdf*) of $f(x,y) = \frac{1}{2}$ for $0 \le x \le y \le 2$ and *f(x,y)*=0 otherwise. Note that the values of X are less than or equal to the values of Y.

4A) **(2 marks)** Verify that the volume under the joint *pdf* equals 1.

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

$$= \frac{1}{2}\left\{\frac{4}{2}\right\} = 1$$

$$\Rightarrow f(x,y) = \frac{1}{2} \text{ is a p.d.f.}$$

4B) (**2 marks**) Find the marginal *pdfs* of X and Y.

$$f_x(x) = \int_x^2 \frac{1}{2}\,dy = \left[\frac{y}{2}\right]_x^2 = \left[1 - \frac{x}{2}\right] \Rightarrow \boxed{f_x(x) = 1 - \frac{x}{2}} \quad 0 \le x \le 2$$

$$f_y(y) = \int_0^y \frac{1}{2}\,dx = \left[\frac{x}{2}\right]_0^y = \frac{y}{2} \Rightarrow f_y(y) = y/2 \quad 0 \le y \le 2$$

4C) (**2 marks**) Find P(X < ½).

$$P(X < \tfrac{1}{2}) = \int_0^{1/2} \left(1 - \frac{x}{2}\right) dx = \left[x - \frac{x^2}{4}\right]_0^{1/2}$$

$$= \frac{1}{2} - 0.0625 = 0.4375$$

4D) (**2 marks**) Find the cumulative density function (*cdf*) of Y.

$$F_y(y) = \int_0^y \frac{y}{2}\,dy = \left[\frac{y^2}{4}\right]_0^y = \frac{y^2}{4}$$

4E) (2 marks) Find the conditional probability P(X<½ |Y=1.5). Are X and Y independent?

first we need to find $f(x|y) = \frac{f(x,y)}{f_y(y)}$

$\Rightarrow f(x|y) = \frac{1/2}{y/2} \Rightarrow f(x|y) = 1/y \quad \because y = 1.5 \Rightarrow f(x|y=1.5) = \frac{2}{3}$

$\Rightarrow P(x<\frac{1}{2}|y=1.5) = \int_0^{\frac{1}{2}} \frac{2}{3} dx = \frac{2}{3} x \Big|_0^{\frac{1}{2}} = \frac{1}{3}$

Note that the $P(x<\frac{1}{2}|y=1.5) \neq P(x<\frac{1}{2})$

$\Rightarrow$ X & Y are not independent

4F) (2 marks) Find the expected value and variance of Y.

$E(Y) = \int_0^2 y\left(\frac{y}{2}\right) dy = \frac{1}{2}\left[\frac{y^3}{3}\right]_0^2 = \frac{1}{2}\left[\frac{8}{3}\right] \Rightarrow E(Y) = \frac{4}{3}$

$E(Y^2) = \int_0^2 \frac{y^3}{2} dy = \frac{1}{2}\left[\frac{y^4}{4}\right]_0^2 = \frac{1}{2}\left[\frac{16}{4}\right] = 2$

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

$= 2 - \frac{16}{9}$

$\Rightarrow Var(Y) = 0.2222$

Q5. **(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 $\tilde{Y} = \frac{1}{2}Y_1 + \frac{1}{3}Y_2 + \frac{1}{6}Y_3$.

5A) **(2 marks)** Prove that $\bar{Y}$ is a linear unbiased estimator.

$\bar{Y} = \frac{1}{3}Y_1 + \frac{1}{3}Y_2 + \frac{1}{3}Y_3 \Rightarrow \bar{Y} = \sum_{i=1}^{3}\left(\frac{1}{3}\right)Y_i$ "linear in Y_i"

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

$= \frac{1}{3}\mu + \frac{1}{3}\mu + \frac{1}{3}\mu \Rightarrow E(\bar{Y}) = \mu \Rightarrow \bar{Y}$ is an unbiased estimator of μ.

5B) **(2 marks)** Prove that $\tilde{Y}$ is a linear unbiased estimator.

$\tilde{Y} = \frac{1}{2}Y_1 + \frac{1}{3}Y_2 + \frac{1}{6}Y_3$, Let $w_1 = \frac{1}{2}$, $w_2 = \frac{1}{3}$, $w_3 = \frac{1}{6} \Rightarrow \tilde{Y} = \sum_{i=1}^{3} w_i Y_i$ "linear in Y_i"

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

$\Rightarrow E(\tilde{Y}) = \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{6}\right)\mu = \mu \Rightarrow \tilde{Y}$ is an unbiased estimator of μ.

5C) **(2 marks)** Calculate the variance of $\bar{Y}$.

$Var(\bar{Y}) = \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2$, Note $Cov(Y_i, Y_j) = 0$ for all $i \neq j$

$Var(\bar{Y}) = \frac{\sigma^2}{3}$

5D) **(2 marks)** Calculate the variance of $\tilde{Y}$.

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

$= \sigma^2\left(\frac{14}{36}\right)$

5E) **(2 marks)** Briefly explain which estimator is better.

From parts A & B, we know that both estimators are unbiased. However, $Var(\bar{Y}) < Var(\tilde{Y})$

$\Rightarrow \bar{Y}$ is a better estimator since it has smaller variance.

Q6. **(12 marks)** The average amount of studying required to get high grades in a course like Statistical Methods II is 6 hours per week outside of class. A professor randomly selects eight students from his class and finds that they are studying 1, 3, 4, 4, 6, 6, 8 and 12 hours, respectively.

6A) **(2 marks)** State the null and alternative hypotheses appropriate with testing whether students are studying an average of six hours per week for Statistical Methods II?

Let μ be the average number of hours that a student devote for studying.

$\Rightarrow$ $H_0: \mu = 6$

$H_1: \mu \neq 6$

we could also use $H_1: \mu > 6$ or $H_1: \mu < 6$

6B) **(2 marks)** Briefly explain whether a *t*- or *z*-statistic is more appropriate.

Assuming that the studying hours are normally distributed with mean (μ) & variance (σ^2), where μ & σ^2 are unknown parameters. Therefore, when testing for $H_0: \mu = 6$ we should use a t-statistic because we do not know (σ^2).

6C) **(2 marks)** Calculate the test statistic.

I first we should estimate μ by computing the sample average: $\bar{Y} = \dfrac{1+3+4+4+6+6+8+12}{8} = 5.5$

II we have to estimate $\hat{\sigma}^2 = \dfrac{\sum(Y_i - \bar{Y})^2}{N-1} = \dfrac{20.25 + 6.25 + 2.25 + 2.25 + \frac{1}{4} + \frac{1}{4} + 6.25 + 42.25}{(8-1)}$

$= \dfrac{80}{7} = 11.42857$

$\Rightarrow \hat{\sigma} = \sqrt{11.42857} = 3.3806$

III Compute the $\widehat{Var}(\bar{Y})$ & $se(\bar{Y})$; Recall $\widehat{Var}(\bar{Y}) = \dfrac{\hat{\sigma}^2}{N} = \dfrac{11.42857}{8}$

$\Rightarrow \widehat{Var}(\bar{Y}) = 1.42857$

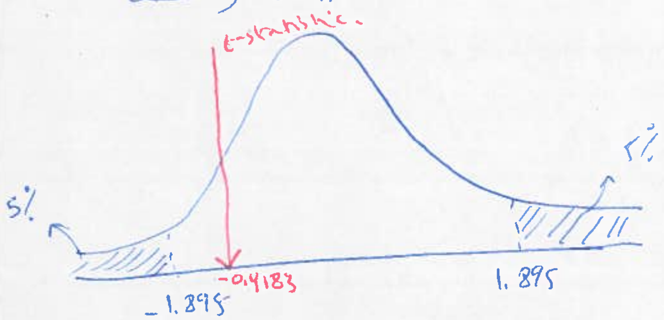
$\boxed{se(\bar{Y}) = \sqrt{1.42857} = 1.19523}$

IV $t = \dfrac{\bar{Y} - 6}{se(\bar{Y})} \Rightarrow t = \dfrac{5.5 - 6}{1.19523}$

$\Rightarrow t = -0.418$

6D) **(2 marks)** Assume that the critical values associated with a 0.10 level of significance are ±1.895. Briefly explain your conclusion. I will solve for $H_0: \mu = 6$, $H_1: \mu \neq 6$

$\Rightarrow$ It is a two tail test



$\therefore -t_c \leq$ t-statistic $\leq +t_c$

$\Rightarrow$ we fail to reject $H_0: \mu = 6$

6E) **(2 marks)** Construct a 90-percent confidence interval.

a 90% interval estimate of μ is $\bar{Y} \pm se(\bar{Y})\, t_c$

$$\Rightarrow \mu \in [5.5 - (1.19523)(1.895), 5.5 + (1.19523)(1.895)]$$

$$\Rightarrow \mu \in [3.235, 7.76496]$$

6F) **(2 marks)** Briefly explain whether your answer in (6E) supports or contradicts your conclusion in (6D).

Yes it does since $\mu = 6$ belongs to the interval estimate we derived in part (E).