MIDTERM EXAMINATION # 1

Statistics 305

Term 1, 2006-2007

Thursday, October 12, 2006

Time: 9:30am – 10:45am

Student Name (**Please print in caps**): SOLUTIONS

Student Number: ______________________

Notes:

- This midterm has 5 problems on the 7 following pages, plus 3 pages of statistical tables. Check to ensure that you have a complete paper.
- The amount each part of each question is worth is shown in [ ] on the left-hand side of the page.
- Where appropriate, record your answers in the blanks provided on the right-hand side of the page.
- Your solutions must be justified; show all the work and state all the reason(s) leading to your answer for each question in the space provided immediately under the question.
- Clear and complete solutions are essential; little partial credit will be given.
- This is a closed book exam.
- A single one-sided 8.5 x 11 page of notes is allowed.
- Calculators are allowed (but not for symbolic differentiation or integration).
- No devices (including calculators) that can store text or send/receive messages are allowed.

Problem	Total Available	Score
1.	12	
2.	5	
3	8	
4.	11	
5.	14	
Total	50	

1. Suppose X and Y are independent and identically distributed random variables that are uniformly distributed on [0, 1].

[2] a) $E(2X-4Y+5) =$ $\underline{\quad 4 \quad}$

$$E(X) = E(Y) = \tfrac{1}{2} \Rightarrow E(2X-4Y+5) = 2(\tfrac{1}{2}) - 4(\tfrac{1}{2}) + 5$$
$$= 1 - 2 + 5$$
$$= \underline{\underline{4}}$$

[3] b) $SD(2X-4Y+5) =$ $\underline{\quad \sqrt{5/3} \quad}$

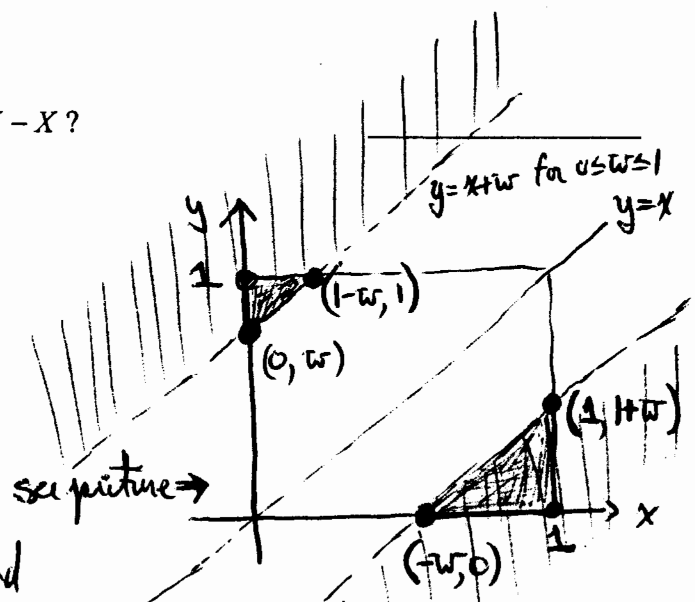
$$Var(2X-4Y+5) = 4\,Var(X) + 16\,Var(Y) \quad \text{since X and Y are independent}$$
$$= 4(\tfrac{1}{12}) + 16(\tfrac{1}{12})$$
$$= \tfrac{1}{3} + \tfrac{4}{3} = \tfrac{5}{3} \Rightarrow \underline{\underline{SD = \sqrt{5/3}}}$$

[7] c) What is the probability density function of $W = Y - X$?

$$F_W(w) = P(W \le w) = 0 \quad \text{for } w < -1$$
$$= 1 \quad \text{for } w > 1$$

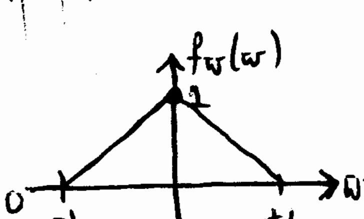
For $-1 \le w \le 1$, need to evaluate

$$P(W \le w) = P(Y - X \le w)$$
$$= P(Y \le w + X) \quad \text{see picture} \rightarrow$$



Because X and Y are independent and both U[0,1] ⇒ probabilities are simply volumes under surface $f_{X,Y}(x,y) \equiv 1$ on top of unit square

$$\Rightarrow P(W \le w) = \tfrac{1}{2}(1+w)^2 \quad \text{for } -1 \le w \le 0$$
$$= 1 - \tfrac{1}{2}(1-w)^2 \quad \text{for } 0 \le w \le 1$$



$$\Rightarrow f_W(w) = \begin{cases} 1+w & \text{for } -1 \le w \le 0 \\ 1-w & \text{for } 0 \le w \le 1 \\ 0 & \text{otherwise} \end{cases}$$

2. Suppose that a measurement has mean μ and standard deviation $\sigma = 2$. We will use $\overline{X}$, the average of n such independent measurements, to estimate the value of μ. How large a value of n is required to be 90% confident that this estimate will be within 0.1 of the true value; that is, $P(|\overline{X} - \mu| \leq 0.1) = 0.90$? 1,083

[5]

By CLT, $\bar{X} \approx N(\mu, \sigma^2/n)$, provided n is "large enough"

$$\Rightarrow P(|\bar{x}-\mu| \leq 0.1) = P\left(\frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}} \leq \frac{0.1}{\sigma/\sqrt{n}}\right)$$

$$\doteq P\left(|Z| \leq \frac{0.1\sqrt{n}}{\sigma}\right) \text{ using CLT}$$

This will be at least 0.90 provided $\frac{0.1\sqrt{n}}{\sigma} \geq 1.645$

$$\Leftrightarrow \sqrt{n} \geq \frac{1.645\,\sigma}{0.1}$$

$$\Leftrightarrow \sqrt{n} \geq 16.45\,\sigma$$

Using $\sigma = 2$ $\Leftrightarrow \sqrt{n} \geq 32.9$

$$\Leftrightarrow n \geq 1{,}082.4$$

3. Suppose $\overline{X}$ and S^2 are the sample mean and the sample variance of a simple random sample of size $n = 9$ from a normal population with mean $\mu = 2$ and variance $\sigma^2 = 36$.

[3] a) $P(0 \le \overline{X} \le 3) =$ 0.53

Sample from normal popln $\Rightarrow \overline{X} \sim N(\mu, \sigma^2/n)$ EXACT!

So, desired probability $= P\left(\frac{0-\mu}{\sigma/\sqrt{n}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{3-\mu}{\sigma/\sqrt{n}} \right)$

But $\frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{9}} = 2 \Rightarrow = P\left(\frac{-2}{2} \le Z \le \frac{1}{2} \right)$

$= P(-1 \le Z \le 1/2) = 0.3413 + 0.1915$

$= \underline{\underline{0.5328}}$

[3] b) $P(15.7 \le S^2 \le 90.4) =$ 0.89

sample from normal $\Rightarrow \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ EXACT!

So, desired probability $= P\left(\frac{(n-1)}{\sigma^2} \times 15.7 \le \frac{(n-1)S^2}{\sigma^2} \le \frac{(n-1)}{\sigma^2} \times 90.4 \right)$

$\frac{(n-1)}{\sigma^2} = \frac{8}{36}$ $= P(3.49 \le \chi^2_{(8)} \le 20.09)$

$= 1 - [0.10 + 0.01] = \underline{\underline{0.89}}$

[2] c) $P(0 \le \overline{X} \le 3 \text{ and } 15.7 \le S^2 \le 90.4) =$ 0.47

Sample from normal $\rightarrow \overline{X}$ and S^2 are independent

So, desired probability $= P(0 \le \overline{X} \le 3) \cdot P(15.7 \le S^2 \le 90.4)$

$= (0.5328)(0.89)$

$= \underline{\underline{0.4742}}$

4. a) Suppose X is normally distributed with mean μ and variance σ^2. Show that
[5] $M_X(t)$, the moment generating function of X, is given by

$$M_X(t) = \exp(\mu t + \tfrac{1}{2}\sigma^2 t^2).$$

Normal density, with mean $= \mu$, var $= \sigma^2$

$$M_X(t) \equiv E(e^{tX}) = \int e^{tx} f_X(x)\,dx$$

$$= \int e^{tx} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} dx$$

$$= \int \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}\left[(x-\mu)^2 - 2\sigma^2 t x\right]\right\} dx$$

Went to "complete the square" in x that appears in the exp term

$$= \int \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}\left[x^2 - 2\mu x + \mu^2 - 2\sigma^2 t x\right]\right\} dx$$

$$= \int \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}\left[x^2 - 2x(\mu + t\sigma^2) + \mu^2\right]\right\} dx$$

$$= \int \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}\left[x^2 - 2x(\mu + t\sigma^2) + (\mu + t\sigma^2)^2 - (\mu + t\sigma^2)^2 + \mu^2\right]\right\} dx$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\left[\mu^2 - (\mu + t\sigma^2)^2\right]\right\} \cdot \int \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2\sigma^2}\left[x - (\mu + t\sigma^2)\right]^2\right\} dx$$

A normal density, with mean $= \mu + t\sigma^2$, var $= \sigma^2$ $\Rightarrow \int \equiv 1$

$$= \exp\left\{-\frac{1}{2\sigma^2}\left[\cancel{\mu^2} - (\cancel{\mu^2} + 2\mu t\sigma^2 + t^2\sigma^4)\right]\right\} \cdot 1$$

$$= \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}$$

[3] b) Show that the relationship between the moment generating function of $\overline{X}$, the mean of a simple random sample of size from any population, and the moment generating function of a single random variable, X, from that same population is given by:

$$M_{\overline{X}}(t) = [M_X(t/n)]^n$$

$$M_{\overline{X}}(t) \equiv E\left(e^{t\overline{X}}\right) = E\left[e^{\frac{t}{n}(X_1+X_2+\cdots+X_n)}\right]$$

$$= \prod_{i=1}^{n} E\left(e^{\frac{t}{n}X_i}\right) \quad \text{because the } X_i\text{'s are independent}$$

$$= \prod_{i=1}^{n} M_{X_i}\left(\frac{t}{n}\right)$$

$$= \left[M_X\left(\frac{t}{n}\right)\right]^n \quad \text{because the } X_i\text{'s all have the \underline{same} distribution}$$

<u>Note</u>: This result has nothing to do with the normal distribution; it is <u>always</u> <u>true</u>!

[3] c) Using the above results, show that the sample mean from a normal population (with mean μ and variance σ^2) is normally distributed with mean μ and variance σ^2/n.

Using b) $M_{\overline{X}}(t) = \left[M_X\left(\frac{t}{n}\right)\right]^n$

From a) $M_X\left(\frac{t}{n}\right) = \exp\left\{\mu\left(\frac{t}{n}\right) + \frac{1}{2}\sigma^2\left(\frac{t}{n}\right)^2\right\}$ for $X \sim N(\mu, \sigma^2)$

$$\Rightarrow M_{\overline{X}}(t) = \left[\exp\left\{\mu\left(\frac{t}{n}\right) + \frac{1}{2}\sigma^2\frac{t^2}{n^2}\right\}\right]^n$$

$$= \exp\left\{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right\}$$

But this is the moment generating function of a normal random variable, with mean $= \mu$ and variance $= \frac{\sigma^2}{n} \Rightarrow \overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

5. Suppose $X_1, X_2, \ldots, X_n$ is a simple random sample from the distribution:

$$f_\theta(x) = \theta x^{\theta-1}. \qquad \text{for } 0 \le x \le 1.$$

Note that this is a density function provided that $\theta > 0$.

$$\frac{\bar{X}}{(1-\bar{X})}$$

[4] a) Find $\hat{\theta}_{MM}$, the method of moments estimator (MME) of θ.

$$\mu_1 \equiv E(X) = \int x f_\theta(x)\,dx$$

$$= \int_0^1 x\,\theta x^{\theta-1}\,dx = \theta \int_0^1 x^\theta\,dx$$

$$= \theta\left[\frac{x^{\theta+1}}{(\theta+1)}\Big|_0^1\right]$$

$$= \frac{\theta}{\theta+1}$$

$$\Rightarrow \theta = \frac{\mu_1}{(1-\mu_1)} \qquad \Rightarrow \hat{\theta}_{MM} = \frac{\hat{\mu}_1}{(1-\hat{\mu}_1)} \equiv \frac{\bar{X}}{(1-\bar{X})}$$

$$\frac{1}{n} \cdot \frac{\theta(\theta+1)^2}{(\theta+2)}$$

[5] b) Find an approximation to the variance of the MME $\hat{\theta}_{MM}$.

$$\hat{\theta}_{MM} = \frac{\bar{X}}{(1-\bar{X})} = g(\bar{X}) \text{ say}$$

Use the delta method to obtain

$$Var(\hat{\theta}_{MM}) \approx \left[g'(E(\bar{X}))\right]^2 \cdot Var(\bar{X})$$

Now $g(x) = x(1-x)^{-1}$

$$\Rightarrow g'(x) = x(-1)(1-x)^{-2}(-1) + (1-x)^{-1}$$

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

From above, $E(X) = \dfrac{\theta}{\theta+1}$

$$\Rightarrow E(\bar{X}) = \frac{\theta}{\theta+1} \Rightarrow g'(E(\bar{X})) = \frac{1}{\left(1-\frac{\theta}{\theta+1}\right)^2} = (\theta+1)^2$$

Also: $Var(\bar{X}) = Var(X)/n$

But $E(X^2) = \dfrac{\theta}{\theta+2}$

$$\Rightarrow Var(X) = \frac{\theta}{\theta+2} - \left(\frac{\theta}{\theta+1}\right)^2 = \frac{\theta}{(\theta+2)(\theta+1)^2}\left[(\theta+1)^2 - \theta(\theta+2)\right]$$

$$= \frac{\theta}{(\theta+2)(\theta+1)^2}$$

Finally, we obtain

$$Var(\hat{\theta}_{mm}) \approx (\theta+1)^4 \frac{\theta}{(\theta+2)(\theta+1)^2} \Big/ n$$

$$\frac{1}{n}\frac{\theta(\theta+1)}{(\theta+2)}$$

[5] c) Find a second-order approximation to the bias of the MME $\hat{\theta}_{MM}$.

Need to evaluate $\text{Bias}(\hat{\theta}_{MM}) = E(\hat{\theta}_{MM}) - \theta$

$\hat{\theta}_{MM} = g(\bar{X})$, where $g(x) = x(1-x)^{-1}$ as in b)

$$\Rightarrow E(\hat{\theta}_{MM}) \approx g[E(\bar{X})] + \frac{1}{2} g''[E(\bar{X})] \cdot \text{Var}(\bar{X})$$

second-order approximation

$$= \theta + \frac{1}{2} \underbrace{\frac{1}{n}\frac{\theta}{(\theta+2)(\theta+1)^2}}_{\text{Var}(\bar{X}) \text{ from b)}} g''(E(\bar{X}))$$

But $g'(x) = (1-x)^{-2}$

$$\Rightarrow g''(x) = -2(1-x)^{-3}(-1)$$

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

$$E(\bar{X}) = \frac{\theta}{\theta+1}$$

$$\Rightarrow g''(E(\bar{X})) = \frac{2}{\left(1-\frac{\theta}{\theta+1}\right)^3} = \frac{2(\theta+1)^3}{1^3} = 2(\theta+1)^3$$

$$\Rightarrow E(\hat{\theta}_{MM}) \approx \theta + \frac{1}{2}\frac{1}{n}\frac{\theta}{(\theta+2)(\theta+1)^2} 2(\theta+1)^3$$

$$= \theta + \frac{1}{n}\frac{\theta(\theta+1)}{(\theta+2)}$$

$$\Rightarrow \text{Bias}(\hat{\theta}_{MM}) \approx \frac{1}{n}\frac{\theta(\theta+1)}{(\theta+2)}$$