MIDTERM EXAMINATION # 1

Statistics 305

Term 1, 2005-2006

Thursday, October 13, 2005

Time: 9:30am – 10:45am

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

Student Number: ______________________________

Notes:

- This midterm has 6 problems on the 6 following pages, plus a final page containing a table of the standard normal distribution. 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 the work and state 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.	7	
2.	7	
3	15	
4.	5	
5.	10	
6.	6	
Total	50	

1. Suppose the moment generating function (mgf) of the random variable X is given by:

$$M_X(t) = [\exp(t) + 2\exp(2t) + 3\exp(3t)] / 6.$$

[2] a) $E(X) =$ $7/3$

$$M_X'(t) = \frac{1}{6}\left[e^t + 2^2 e^{2t} + 3^2 e^{3t}\right]$$

$$\Rightarrow M_X'(0) = \frac{1}{6}\left[1+4+9\right] = \frac{14}{6}$$

[3] b) $SD(X) =$ $\sqrt{5/9}$

$$M_X''(t) = \frac{1}{6}\left[e^t + 2^3 \cdot e^{2t} + 3^3 \cdot e^{3t}\right]$$

$$\Rightarrow M_X''(0) = \frac{1}{6}\left[1+8+27\right] = \frac{36}{6} = 6$$

$$\Rightarrow Var(X) = 6 - (7/3)^2 = 6 - \frac{49}{9} = 6 - 5\tfrac{4}{9} = 5/9$$

Alternately, could express $M_X(t)$ as a power series in t to solve a) and b):

$$M_X(t) = \frac{1}{6}\left[\sum_{k=0}^{\infty}\frac{t^k}{k!} + 2\sum_{k=0}^{\infty}\frac{(2t)^k}{k!} + 3\sum_{k=0}^{\infty}\frac{(3t)^k}{k!}\right]$$

$$= \frac{1}{6}\left[\sum_{k=0}^{\infty}(1 + 2\cdot 2^k + 3\cdot 3^k)\, t^k/k!\right]$$

This yields $E(X) = (1 + 2\cdot 2 + 3\cdot 3)/6$, and $E(X^2) = (1 + 2\cdot 2^2 + 3\cdot 3^2)/6$ etc.

[2] c) What is the distribution of the random variable X?

$$X = \begin{cases} 1 & \text{w.p. } 1/6 \\ 2 & \text{w.p. } 2/6 \\ 3 & \text{w.p. } 3/6 \end{cases}$$

Consider the discrete r.v. Y that is 1 w.p. $1/6$, 2 w.p. $2/6$ and 3 w.p. $3/6$

$$\Rightarrow M_Y(t) = \frac{1}{6}\left[e^t + 2e^{2t} + 3e^{3t}\right] \equiv M_X(t)$$

As MGFs uniquely determine distributions $\Rightarrow$ the distribution of X is this <u>same</u> distribution!

2. Suppose X is a binomial random variable with parameters n and p.

[2] a) Show that $M_X(t)$, the mgf of X, is given by

$$M_X(t) = [1 - p + p\exp(t)]^n.$$

For what values of t does $M_X(t)$ exist?

The p.m.f of X is given by $f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, 2, \cdots, n$

$$\Rightarrow E(e^{tX}) = \sum e^{tx} f_X(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} [pe^t]^x (1-p)^{n-x}$$

$$\Rightarrow = [1 - p + pe^t]^n$$

Exists for any value of t

$$\boxed{\text{Binomial expansion: } (a+b)^n \equiv \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}}$$

[3] b) Show that if $n \to \infty$ and $p \to 0$ in such a way that $np \to \lambda$, then

$$M_X(t) \to \exp\{\lambda [\exp(t) - 1]\}.$$

Set $\lambda = np + \epsilon$, where $\epsilon \to 0$ as $n \to \infty$

$$M_X(t) = [1 + p(e^t - 1)]^n \overset{\Downarrow}{=} \left[1 + \left(\frac{\lambda + \epsilon}{n}\right)(e^t - 1)\right]^n$$

$$= \left[1 + \frac{\lambda}{n}(e^t - 1) + \frac{\epsilon}{n}(e^t - 1)\right]^n$$

$$\to e^{\lambda(e^t - 1)}$$

because $\epsilon \to 0$ as $n \to \infty$

$$\boxed{\text{But, if } a_n \to a \text{ as } n \to \infty \text{ then } \left(1 + \frac{a_n}{n}\right)^n \to e^a \text{ as } n \to \infty}$$

[2] c) Explain the practical importance of the result in b).

Limit is the MGF of a Poisson r.v. with parameter λ

$\Rightarrow$ Distribution of the binomial r.v. (with parameters n and p) approaches that of Poisson (with parameter λ) in this limit

$\Rightarrow$ For large n and small p, can <u>approximate</u> binomial probabilities by Poisson probabilities (with $\lambda = np$)!

3. Suppose X, the lifetime for an electronic device, is an exponential random variable with a mean of 2 years; that is, $f_X(x) = \frac{1}{2}\exp(-x/2)$ for $x > 0$. Suppose 100 of these electronic devices are operating independently.

Provide an accurate numerical approximation to the probability that:

[6] a) at least 50 of these 100 devices are still operating after 2 years. 0.0042

$$P(X>2) = \int_2^\infty \tfrac{1}{2}\exp(-x/2)\,dx = e^{-1} \cong 0.3679$$

If N = # out of the 100 that are still operating after 2 years,

then $N \sim B(100, e^{-1})$ ← because of independence

$$P(N \geq 50) = \sum_{x=50}^{100} \binom{100}{x}(e^{-1})^x[1-e^{-1}]^{100-x} = P(N \geq 49.5) = P\left(\frac{N-100e^{-1}}{\sqrt{100e^{-1}(1-e^{-1})}} \geq \frac{49.5-100e^{-1}}{\sqrt{}}\right)$$

$$\cong P\left(Z \geq \frac{12.7121}{\sqrt{23.2544}}\right)$$

$$= P(Z \geq 2.6361) \cong 0.0042$$

↑ exact

Using R, the exact value is 0.0047

If you don't use continuity correction, you will get $\cong P(Z \geq 2.7390) \cong 0.0031$

[5] b) less than 3 of these 100 devices are still operating after 8 years. 0.72

$$P(X>8) = \int_8^\infty \tfrac{1}{2}\exp(-x/2)\,dx = e^{-4} \cong 0.0183$$

If N^* = # out of 100 still operating after 8 years,

then $N^* \sim B(100, e^{-4})$

From #2, b) and c)

$$P(N^* < 3) = P(N^* \leq 2) = \sum_{x=0}^{2}\binom{100}{x}(e^{-4})^x[1-e^{-4}]^{100-x} \cong P(N' \leq 2) \text{ where } N' \sim P(100e^{-4})$$

$$= e^{-[100e^{-4}]}\left(1 + \frac{[100e^{-4}]}{1!} + \frac{[100e^{-4}]^2}{2!}\right)$$

$$= 0.1602(4.5089) \cong 0.7222$$

↑ exact

Exact value is not so hard to get in this case $\cong 0.7226$, but I asked for "accurate numerical approx"

[4] c) the average value of the 100 lifetimes is at least 2.3 years. 0.067

$E(X) = 2$

$Var(X) = 4 \Rightarrow SD(X) = 2$ $\Rightarrow \bar{X} \approx N\left(2, \frac{2^2}{100}\right)$ by CLT

$$P(\bar{X} \geq 2.3) = P\left(\frac{\bar{X}-2}{2/10} \geq \frac{2.3-2}{2/10}\right)$$

$$\cong P(Z \geq 1.5) \cong 0.0668$$

4. Suppose X is a positive random variable with mean μ_X and variance σ_X^2. If $Y = \ln(X)$, where $\ln = \log_e$, use the delta method to obtain expressions (in terms of μ_X and σ_X) for:

[1] a) a "first-order" approximation to E(Y). $\underline{\log \mu_X}$

In general: $f(X) = f(\mu_X) + (X-\mu_X) f'(\mu_X) + \frac{1}{2}(X-\mu_X)^2 f''(\mu_X) + \cdots$

$\Rightarrow$ To first-order: $f(X) \cong f(\mu_X) + (X-\mu_X) f'(\mu_X)$

Note: Need X positive for this to make any sense!

$\Rightarrow E[f(X)] \cong f(\mu_X) \Rightarrow \log \mu_X$ for $f(x) = \log x$

[2] b) a "first-order" approximation to SD(Y). $\underline{\sigma_X / \mu_X}$

From above, to first-order, we have:

$$\mathrm{Var}[f(X)] \cong [f'(\mu_X)]^2 \cdot \mathrm{Var}(X)$$

$$= \left(\frac{1}{\mu_X}\right)^2 \sigma_X^2 = \sigma_X^2 / \mu_X^2$$

$$\Rightarrow SD[f(X)] \cong \sigma_X / \mu_X$$

[2] c) a "second-order" approximation to E(Y). $\underline{\log \mu_X - \frac{1}{2}\left[\frac{\sigma_X}{\mu_X}\right]^2}$

From above, to second-order, we have

$$E[f(X)] \cong f(\mu_X) + \frac{1}{2} f''(\mu_X) \cdot \mathrm{Var}(X)$$

$\Rightarrow$ For $f(x) = \log x$, this becomes $\log \mu_X + \frac{1}{2}\left[-\frac{1}{\mu_X^2}\right]\sigma_X^2$

$$= \log \mu_X - \frac{1}{2} \sigma_X^2 / \mu_X^2$$

5. Suppose X and Y are bivariate normally distributed random variables with means μ_X, μ_Y and variances σ_X^2, σ_Y^2, respectively, and correlation ρ. Then, as given in class, $M_{X,Y}(s,t)$, the joint mgf of X and Y evaluated at s and t, is given by:

$$M_{X,Y}(s,t) = \exp\{\, s\,\mu_X + t\,\mu_Y + [\, s^2\,\sigma_X^2 + 2\,st\,\rho\,\sigma_X\,\sigma_Y + t^2\,\sigma_Y^2\,]\,/2\,\}.$$

Let $W = aX + bY + c$, where a, b and c are constants.

a) Give the expressions for:

[1] i) E(W)

[2] ii) Var(W)

These are immediate — don't need to know anything about bivariate normal for this as expressions are true in general

$a\mu_X + b\mu_Y + c$

$a^2\sigma_X^2 + 2ab\rho\sigma_X\sigma_Y + b^2\sigma_Y^2$

[5] b) Evaluate the mgf of W.

Below

$$M_W(t) = E(e^{tW}) = E\left(e^{t[aX+bY+c]}\right)$$
$$= e^{tc}\, E\left[e^{(at)X+(bt)Y}\right]$$
$$= e^{tc}\, M_{X,Y}(at, bt)$$
$$= e^{tc} \exp\left\{(at)\mu_X + (bt)\mu_Y + \left[(at)^2\sigma_X^2 + 2(at)(bt)\rho\sigma_X\sigma_Y + (bt)^2\sigma_Y^2\right]/2\right\}$$
$$= \exp\left\{[a\mu_X + b\mu_Y + c]\,t + \left[a^2\sigma_X^2 + 2ab\rho\sigma_X\sigma_Y + b^2\sigma_Y^2\right] t^2/2\right\}$$

[2] c) What is the distribution of W?

$N(\mu, \sigma^2)$, where

$$M_W(t) = \exp\{\mu t + \tfrac{1}{2}\sigma^2 t^2\}, \text{ where}$$

Of course, you should have guessed that this would be the answer!

$$\left.\begin{aligned} \mu &= a\mu_X + b\mu_Y + c \\ \sigma^2 &= a^2\sigma_X^2 + 2ab\rho\sigma_X\sigma_Y + b^2\sigma_Y^2 \end{aligned}\right\}$$

$$\Rightarrow W \sim N(\mu, \sigma^2)$$

Linear combinations of independent normal r.v.'s are normally distributed

6. Suppose X and Y are independent normally distributed random variables with:

$$\mu_X = 8,\ \sigma_X^2 = 9 \quad \text{and} \quad \mu_Y = 3,\ \sigma_Y^2 = 16.$$

[6] Evaluate P($X > Y$).

0.84

As $X-Y \sim N(5,\ 9+16)$, $\Leftarrow$ by independence

[or, use the more general result in #5 c)]

we can evaluate

$$P(X>Y) = P(X-Y>0)$$

$$= P\left(\frac{X-Y-5}{\sqrt{9+16}} > \frac{-5}{5}\right)$$

$$= P(Z > -1) \cong 0.8413$$