University of Ottawa Faculty of Administration

ADM 2303: STATISTICS FOR MANAGEMENT I
SPECIAL FINAL EXAMINATION January 2000

NAME: S.N. Section: A B C D

Time: 3 hours Total marks:60

ALL ANSWERS (INCLUDING BRIEF EXPLANATIONS) GO ON THE ANSWER SHEET. THE EXAM QUESTION SHEETS WILL **NOT** BE MARKED, though space is provided here for your rough work. The question sheets **must** be deposited in the box provided. NOTE THAT THERE ARE MARKS FOR EXPLAINING YOUR ANSWERS, SO MAKE SURE YOU INCLUDE BRIEF EXPLANATIONS ON THE **ANSWER** SHEET. THERE ARE MARKS FOR IDENTIFYING PROBABILITY DISTRIBUTIONS. Calculators, 1 sheet of notes, on 8.5" by 11" paper (no stick-ons!). You do **not** need to interpolate, but take the nearest table value.

Q 1. The Ottawa area now has approximately 340,000 households. About 0.0082% of these have annual incomes exceeding $10,000,000. On the other hand, about 27.3% households have income exceeding $48,000 per year. Suppose Statistics Canada sends out a "long" census form that asks about income to 1 household in 10 in Ottawa. Other forms do not ask about income, so we are only interested in the "long" form.

a) [3] What is the probability no returned form has an income exceeding $10,000,000?

b) [3] What is the probability 2 forms or more are returned with income exceeding $10,000,000?

A statistician pulls 10 returned forms at random from those returned.

c) [3] What is the probability he finds none with income > $48,000?.

d) [3] What is the probability he finds 3 forms with income > $48,000?

} long forms assumed

e) [5] The statistician now expands his/her checks and pulls 300 forms at random. What is the probability more than 92 have income >$48,000?

f) [4] The statistician has 11 forms reporting incomes exceeding $1,000,000 on his/her desk. Another employee, in violation of the Statistics Act, manages to sneak a look at 5 of the forms in order to find out who are the single millionnaires in order to follow the advice of the film "How to marry a millionnaire". If 4 of the 11 forms are from single millionnaires, what is the chance the sneak learns about at least 2 single millionnaires?

Q 2. Power glitches are a serious problem for computer tape backup. The backup tape drive streams data onto tape at 1.25 Megabytes (MB) per second. A power glitch means we will likely need to start over. (We discover this by doing a file compare of tape to our disk after writing the tape.) Glitches occur randomly at any time at a rate of 2.1 per 24 hours.

a) [5] You need to back up 6.5 Gigabytes of data. (1GB = 1000 MB). What is the probability of no glitches during the backup?

b) [3] The boss now wants you to back up the whole office -- 190 GB. You plan to do this on a set of tapes, each of which holds 10 GB of data and takes 135 minutes to write (we include some setup time here so that the numbers do NOT perfectly match with the rate given above). What is the probability of no glitch on a single tape?

c) [3] Recall that tape holds 10 GB. You back up the computers in the whole office (190 GB) onto tapes. Each tape can be treated independently, so if one has a glitch, this does not affect the next. What probability distribution do you use to compute the probability no tape has been affected by a power glitch? What are its parameters?

d) [4] What is the probability that there would be more than two power glitches in a 48 hour period?

Q 3. You run a small fabrication company that works with exotic metal alloys. You are asked to make a titanium alloy shaft that will rotate in a high-density plastic sleeve. The diameter of the shaft is 10mm. You will measure to millionths of a metre. To allow the spindle to rotate, but not have too much slack, the buyer wants the diameters to be between 9.992 and 9.995 mm.

a) [2] You decide to test the distribution of ~~concentration of the compound~~ diameters. Some graphs for 200 samples are presented, measured in units of 0.001 mm. Based on these, can you say that the SHAPE of the distribution is Gaussian?

Descriptive Statistics

```
Variable      N    Mean    Median   TrMean    StDev   SE Mean
spindiam    200  9985.6    9984.5   9985.3     9.3       0.7

Variable        Minimum    Maximum         Q1         Q3
spindiam         9965.5    10007.8     9978.3     9991.7
```

```
Histogram of spindiam   N = 200

Midpoint  Count
 9965       2  **
 9970       6  ******
 9975      37  *************************************
 9980      44  ********************************************
 9985      31  *******************************
 9990      32  ********************************
 9995      22  **********************
10000      14  **************
10005      11  ***********
10010       1  *
```

```
MTB > boxp c1
```

Boxplot

```
                            ------------------
         ----------------I        +         I--------------------
                            ------------------
      ------+---------+---------+---------+---------+---------+spindiam
        9968.0    9976.0    9984.0    9992.0   10000.0   10008.0
```

b) [4] Regardless of your answer in (a), for this part of the question assume that you have decided that the data is NOT Gaussian and that we will treat it as a sample of 200 observations. What is the probability that this sample would have a mean that is at least the observed mean diameter if it came from a parent population having mean diameter of 9.987 mm and standard deviation of 0.01mm. Be sure to state any assumptions and distributions used.

c) [6] Regardless of your answers in (a) and (b), you decide that the data can be accepted as Gaussian. A further 7 spindles are selected at random and measured. Their diameter in units of 0.001 mm are

10001.7 10000.4 10002.8 9997.9 9993.5 9999.4 9994.6

Would you accept that this sample is drawn from a population having a mean diameter of 9.994 mm? Support your answer with an appropriate calculation of the approximate probability that you could observe such a sample. Be sure to explain any assumptions or theory you use.

d) [3] How would you modify your answer in (c) if the population standard deviation were given as 0.0035 mm? Explain any theory or assumptions you use.

Q4. Regardless of your answers in Q3, you decide to use 9.994 mm and 0.004mm as the specifications for the mean and standard deviation of the diameter specification for the spindles. You also assume an approximate Gaussian distribution of spindle diameter. [For the rest of the question you may use this information.]

a) [4] Approximately what proportion of individual spindles will be ***out of specification***, that is, not having diameters to be between 9.992 and 9.995 mm. Justify BRIEFLY.

b) [2] You decide to check the variability of spindle diameter and take samples of ten measurements at regular time intervals from the production line. This yields the chart:

```
MTB > schart c4 10;
SUBC> sigma=4.
                    S Chart for spindiam
         -
         -
S        -----------------------------------------UCL=6.677
a        -           +
m    6.00+
p        -   +
l        -             +             +     +
e        -       +                           + +
         -                             +
S    4.00+-------------------+-+------------------S=3.891
t        - +               +     +       +
D        -     +   +     +         +             +
e        -
v        -
     2.00+
         -
         -----------------------------------------LCL=1.104
         -
         +---------+---------+---------+---------+
         0         5        10        15        20
                        Sample Number
```

What is your conclusion about the variability of the process of spindle manufacture?

c) [3] Regardless of your answer in b), assume that spindle manufacture goes ahead. What do you conclude from the following chart? You may wish to compare with the graphs in Q3 that use the same data.

```
MTB > xbarchart c1 10;
SUBC> sigma=4;
SUBC> mu=9994.


                  X-bar Chart for spindiam
         -
         -                                       +
         -
S 10000.0+
a        --+--------------------------------------UCL=9998
m        -                                     +
p        ----+-------------------------------+----MU=9994
l        -                                 +
e  9990.0+----------------------------------------LCL=9990
         -     + +
M        -                               +
e        -         +                 + +
a        -           +
n  9980.0+               +
         -             +       + + +
         -
         -                 + +
         -
         +---------+---------+---------+---------+
         0         5        10        15        20
                        Sample Number
```

ADM 2303 - STATISTICS FOR MANAGEMENT I January 2000

NAME: S.N. Section: A B C D

1. a) [3] | b) [3]

---

c) [3] | d) [3]

---

e) [5] | f) [4]

---

2 a) [5] | b) [3]

c2) [3]

d) [4]

Q 3. a) [2]

b) [4]

---

c) [6]

d) [3]

---

4 a) [4]

```
1. a) [ 3 ]
Situation binomial, n=340000/10=34000
 p=0.000082 (= 0.0082%)
P(0 cases | n, p) = (1-p)^n
  = 0.999918^34000 = 0.061537

or use Poisson approximation (MUST mention)
 P(0 | n,p) ~= P_Poisson( 0 |n*p=mu=2.788)
  = exp(-2.788) = 0.061544

  1 mark binomial (actually hypergeom)
  1 mark n, p or other setup
  1 mark answer
```

```
b) [ 3 ]
P(>=2 forms) = 1 - P(K<=1 cases)
 Can compute with binomial or Poisson approx.
 P(>=2) = 1 - (1-p)^n - n p (1-p)^(n-1)
 = 1 - 0.061544 - 0.171571 = 0.766885

  1 mark setup, 1 working, 1 answer

 Could use Poisson approx
 1 - exp(-2.788)(1+2.788) = 0.766871
```

```
c) [ 3 ]
Binomial situation, n = 10, p=0.273
 P(0 >48K$ | 10, p=0.273) = 0.727 ^ 10
    = 0.041242
  1 mark situation, 1 working, 1 answer
```

```
d) [ 3 ]
 Binomial as in (c)
 P(3 forms | 10 tested, p=0.273)
 = C(10,3) * .273^3 *  .727^7
 = 120 * .020346 * .107335 = .262066
  1 setup, 1 working, 1 answer
```

```
e) [ 5 ]
 Binomial situation  n=300, p=0.273
 P(> 92 have >48K$ | n=300, p=0.273)
 ~= P_Gaussian(z > (92.5 - 81.9)/7.7163 )
  = P(z>1.373715) ~= P(z>1.37) = 0.5 - .4147
   =  .0853
  1 mark setup,
  1 mark mu=n*p=81.9, sqrt(n*p*q)=sigma
   =sqrt(59.54) = 7.7163
  1 mark for correction for continuity
  1 mark for z and table use
  1 mark for answer
```

```
f) [ 4 ]
 Hypergeometric  N=11, S=4, n=5, k>=2
 P(0 | N=11, S=4, n=5)
 = C(4,0) * C(7,5)/C(11,5) = 1 * 21 / 462
 = 0.045455
 P(1 | N=11, S=4, n=5)
 = C(4,1) *(C(7,4)/C(11,5) = 4 * 35/463
 = 0.30303
 P(k >=2) = 1 - P(0) - P(1) = 0.651515

  1 setup problem
  1 setup formula
  1 working
  1 answer
```

```
2 a) [ 5 ]
  Poisson situation
  Need rate of glitches = 2.1/24 hrs
    = 2.1/(24*60*60) = 2.43056E-5 per sec.
 Time to backup 6.5 GB = 6500/1.25 secs
 = 5200 secs.
 Expected glitches = 0.126389
 P(0|0.126389 expected) = exp(=0.126389)
 = 0.881272
 1 mark for Poisson
```

```
b) [ 3 ]
 mu = 135*2.1/(24*60) = .19688 expected
 glitches per tape. Still Poisson
 P(0 | 0.19688) = exp(-0.19688) = .82129

 1 mark Poisson, 1 work, 1 answer
```

2 c) [3]
Need 190/10 = 19 tapes
There is a probability of 0.82129 that a tape will NOT have a glitch, and tapes are independent. So we have a BINOMIAL situation with n=19 and p as above.
1 mark binomial, 1 n, 1 p
Note that p just has to agree with (b) does not have to be correct.
1 mark Poisson, 1 work, 1 answer

d) [4]
Clearly Poisson (1 mark)

```
P_Poisson(>2 glitches | 2.1*48/24 expected)
= 1 - P(0) - P(1) - P(2)
= 1 - exp(-4.2)*(1 + 4.2 + 4.2*4.2/2)
= 1 - 0.015 * ( 1+ 4.2+8.82)
= 1 = 0.015*(14.02) = 1 - 0.21024 = 0.78976
```

1 mark Poisson, 1 for mu,
1 for 1 - P(0)-P()1)-P(2), 1 answer

Q 3. a) [2]
While the histogram shows some skewness the boxplot shows a more or less symmetric graph. We therefore accept the data as coming from a Gaussian distribution

Must mention both graphs for full marks.

b) [4]
Data assumed drawn from a distribution with mean 9.987mm and popn s.d. 0.01 mm, NOT assumed Gaussian. Sample n=200 is large enough that we expect Central Limit Theorem to apply.

```
P(Xbar >9.9856 | mu=9.987,sigma(X)=0.01)
~= P(z>=(9.9856-9.987)/(.01/sqrt(200)))
= P(z>-1.9799) ~=P(z>-1.98)=.5+.4761
= 0.9761
```

or almost 98% chance of exceeding observed mean if stated population mean and s.d. true
1 mark setup
1 mark explicit CLT statement
1 mark working
1 mark answer

---

c) [6]
Gaussian parent distn with mu=9.994 mm, sigma=??
Thus must use Student t distribution
Calculate Xbar=9.9986 mm, s=.003506mm

```
P(Xbar>=9.9986|mu=9.994)
= P(t>=(X_bar=mu)/(s/sqrt(7)))
= P(t>=3.48195)
Compare t(6,0.005)=3.707, t(6,.01)=3.143
so P(t>=3.482) is <1% but >0.5%
```

This is a small probability so would NOT accept sample is drawn from stated population.
1 mark justifying t
1 mark Xbar, 1 mark s
1 mark setup t
1 mark table value interpretation
1 mark conclusion/justification

d) [3]
sigma = 0.0035 mm Now know that Xbar is distributed EXACTLY as Gaussian with mu(Xbar)=9.984, sigma(Xbar)=.0035/sqrt(7)

```
P(Xbar>=9.9986)~= P(z>=3.48)
= 0.5 - 0.4997 = .0003
```

Very tiny chance.
1 mark Gaussian
1 mark work
1 mark answer

---

4 a) [4]

```
P(X <= 9.992 mm OR X > 9.995 mm | mu=9.994, mu=0.004) = 1 - P(9.992<X<9.995)
    = 1 - P((9.992-9.994)/0.004 <= z <= (9.995-9.994)/0.004))
    = 1 - P(-0.5 < z < 0.25) = 1 - .1915 - .0987 = 0.7098 or almost 71% chance OUT of spec.
```

1 mark setup, 2 marks work, 1 mark answer