Suggested Exercises from M&M Chapter 5 *Homegrown exercises begin on page 2*

These pages were updated on September 23

To start with, do some of the odd-numbered exercises. answers to all odd-numbered exercises are given on textbook pages S-1 onwards.

Do some or all of the following even-numbered exercises. You are asked to hand in answers to designated ones.. see the list, and the deadline, on the main course page. Some of these will be discussed in tutorials or answers to them posted on the course web page

§ 5.1	§ 5.2
5.2	5.24
5.4*	5.26
5.6	5.29
5.8	5.30
5.10	5.34
5.12	5.40
5.15	5.44
5.16	5.46
5.17	5.74
5.22	

* For 5.4 (a) see comment at top of p 378.

"Homegrown" Exercises around M&M Chapter 5

-1- Clusters of Miscarriages [*based on article by L Abenhaim*]

Assume that:

a 15% of all pregnancies end in a recognized spontaneous abortion (miscarriage)

b across North America, there are 1,000 large companies. In each of them, 10 females who work all day with computer terminals become pregnant within the course of a year [the number who get pregnant would vary, but assume for the sake of this exercise that it is exactly 10 in each company].

c there is no relationship between working with computers and the risk of miscarriage.

d a "cluster" of miscarriages is defined as "at least 5 of 10 females in the same company suffering a miscarriage within a year".

Calculate the number of "clusters" of miscarriages one would expect in the 1,000 companies. Hint: begin with the *probability* of a cluster.

-2- "Prone-ness" to Miscarriages ?

Some studies suggest that the chance of a pregnancy ending in a spontaneous abortion is approximately 30%.

a On this basis, if a woman becomes pregnant 4 times, what does the binomial distribution give as her chance of having 0,1,2,3 or 4 spontaneous abortions?

b On this basis, if 70 women each become pregnant 4 times, what number of them would you expect to suffer 0,1,2,3 or 4 spontaneous abortions? (Think of the answers in a as proportions of women)

c Compare these theoretically expected numbers out of 70 with the following observed data on 70 women:

	No. of spontaneous abortions per 4 pregnancies				
	0	1	2	3	4
No. of women	23	28	7	6	6

d Why don't the expected numbers agree very well with the observed numbers? i.e. which assumption(s) of the Binomial Distribution are possibly being violated? (Note that the overall rate of spontaneous abortions in the observed data is in fact 84 out of 280 pregnancies or 30%)

-3- Automated Chemistries (from Ingelfinger et al)

At the Beth Israel Hospital in Boston, an automated clinical chemistry analyzer is used to give 18 routinely ordered chemical determinations on one order (glucose, BUN, creatinine, ..., iron). The "normal" values for these 18 tests were established by the concentrations of these chemicals in the sera of a large sample of healthy volunteers. The normal range was defined so that an average of 3% of the values found in these healthy subjects fell outside.

a Using the binomial formula, compute the probability that a healthy subject will have normal values on all 18 tests. Also calculate the probability of 2 or more abnormal values.

b Which of the requirements for the binomial distribution are definitely satisfied, and which ones may not be?

c Among 82 normal employees at the hospital, 52/82 (64%) had all normal tests, 19/82 (23%) had 1 abnormal test and 11/82 (13%) had 2 or more abnormal tests. Compare these observed percentages with the theoretical distribution obtained from calculations using the binomial distribution. Comment on the closeness of the fit.

-4- Binomial or Opportunistic (Capitalization on chance.. multiple looks at data) ?"

(from Ingelfinger et al)

Mrs A has mild diabetes controlled by diet. Her morning urine sugar test is negative 80% of the time and positive (+) 20% of the time [It is never graded higher than +].

a At her regular visit she tells her physician that the test has been + on each of the last 5 days. What is the probability that this would occur if her condition has remained unchanged? Does this observation give reason to think that her condition **has** changed?

b Is the situation different if she observes, between visits, that the test is positive on 5 successive days and phones to express her concern?

- 5 - "???" (from Earlier version of M&M)

In 50% of cases of breast cancer, the disease is "localized" at time of diagnosis and of these cases 90% survive 5 years. Of the remainder with "extended" disease only 40% survive 5 years.

a Given that a patient survives 5 years, what is the probability that her disease was localized at time of diagnosis?

b If 5 women are diagnosed as having "localized" breast cancer, calculate the probability that 0, 1, ..., 5 of them will survive 5 years.

- 6 - Can one influence sex of baby?

(NEJM300:1445-1448, 1979)

Consider a binomial variable with n = 145 and $\pi = 0.528$. Calculate the SD of, and therefore a measure of the variation in, the proportions that one would observe in different samples of 145 if $\pi = 0.528$.

Then consider the following, abstracted from NEJM 300: 1445-1448, 1979: and answer the question that follows the excerpt.

> *"The baby's sex was studied in births to Jewish women who observed the orthodox ritual of sexual separation each month and who resumed intercourse within two days of ovulation. The proportion of male babies was 95/145 or 65.5% (!!) in the offspring of those women who resumed intercourse two days after ovulation (the overall percentage of male babies born to the 3658 women who had resumed intercourse within two days of ovulation [i.e. days -2, -1, 0, 1 and 2] was 52.8%)".*

How does the SD you calculated above help you judge the findings?

- 7 - Gaussian Approximation to Binomial: how good is it? (see A&B's table 2.6, p 70)

Compare exact and approximated probabilities for the ≤ 0 and ≥ 4 tails of the Binomial when n =10, $\pi = 0.2$. Comment.

-8- It's the 3rd Tuesday, it must be Binomial !!

In which of the following would X not have a Binomial distribution? Why?

a. The pool of potential jurors for a murder case contains 100 persons chosen at random from the adult residents of a large city. Each person in the pool is asked whether he or she opposes the death penalty; X is the number who say "Yes"

b. Y = number of women listed in different random samples of size 20 from the 1990 directory of statisticians.

c. Y = number of occasions, out of a randomly selected sample of 100 occasions during the year, in which you were indoors. (One might use this design to estimate what proportion of time you spend indoors)

d. Y = number of months of the year in which it snows in Montréal.

e. Y = #, out of 60 occupants of 30 randomly chosen cars, wearing seatbelts.

f. Y = #, out of 60 occupants of 60 randomly chosen cars, wearing seatbelts.

g. Y = #, out of a department's 10 microcomputers and 4 printers, that are going to fail in their first year.

h. Y = #, out of simple random sample of 100 individuals, that are left-handed.

i. Y = #, out of 5000 randomly selected from mothers giving birth each month in Quebec, who will test HIV positive.

j. You observe the sex of the next 50 children born at a local hospital; Y is the number of girls among them.

k. A couple decides to continue to have children until their first girl is born; Y is the total number of children the couple has.

l. You want to know what percent of married people believe that mothers of young children should not be employed outside the home. You plan to interview 50 people, and for the sake of convenience you decide to interview both the husband and the wife in 25 married couples. The random variable Y is the number among the 50 persons interviewed who think mothers should not be employed.

-9- Tests of intuition

a A coin will be tossed either 2 times or 20 times. You will win $2.00 if the number of heads is equal to the number of tails, no more and no less. Which is correct? (i) 2 tosses is better. (ii) 100 tosses is better. (iii) Both offer the same chance of winning. Hint: use Table C.

b Hospital A has 100 births a year, hospital B has 2500. In which hospital is it more that at least 55% of births in one year will be boys.

-10- Aggregation ..

On the average, hotel guests weigh about 150 lb with SD=25 lb. An engineer is designing a large elevator for a convention hotel, to lift 100 people. If he designs it to lift 15,500 lb, the chance it will be overloaded by a random group of 100 people is closest to 0.1 of 1%, 2%, 5%, 50%, 95%, 98%, 99.9%

-11- Births after The Great Blackout

[from Statistics text by Freedman et al.]

On November 9, 1965, the electric power went out in New York City, and it stayed out for a day -- The Great Blackout. Nine months later, newspapers suggested that New York was experiencing a baby boom. The table shows the number of babies born every day during a twenty-five day period, centered nine months and ten days after The Great Blackout. These numbers average out to 436. This turns out to be not unusually high for New York. But there is an interesting twist: the 3 Sundays only average 357.

a How likely is it that the average of three days chosen at random from the table will be 357 or less? What do you infer?

Number of births in New York,
Monday August 1-Thursday August 25, 1966.

Mon	Tue	Wed	Thu	Fri	Sat	Sun
451	468	429	448	466	377	344
448	438	455	468	462	405	377
451	497	458	429	434	410	351
467	508	432	426			

Hint: The SD of the 25 numbers in the table is about 40. Formulate the null hypothesis; the normal approximation can be used.

This question and the following footnote come from the Statistics text by Freedman et al.

> "Apparently, the *New York Times* sent a reporter around to a few hospitals on Monday August 8, and Tuesday August 9, nine months after the blackout. The hospitals reported that their obstetric wards were busier than usual -- apparently because of the general pattern that weekends are slow, Mondays and Tuesdays are busy. These "findings" were published in a front-page article on Wednesday, August 10, 1966, under the headline "Births Up 9 Months After the Blackout." This seems to be the origin of the baby-boom myth."

b Suggest a better plan for estimating the impact , if any, of the Blackout on the number of births.

c (Still on the subject of births, but now in Québec). In an effort to bolster sagging birth rate, the Québec government in it's budget of March 1988 implemented a cash bonus of $4,500 to parents who had a third child. Suggest a method of measuring the impact of this incentive scheme -- be both precise and concise.

-12- Why and when SD? [from previous exams]

Check all correct lettered answers. (None, one, or more than one may be correct): Results are often reported in the medical literature as the sample Mean ± one Standard Deviation (S.D.). It is a good procedure to report the S.D. because:

a It is a measure of variability of individuals.

b It is used to determine the Standard Error of the Mean.

c It influences the 95% confidence limits on the true mean.

d It varies systematically with sample size.

-13- Overlapping Confidence Intervals

Later on, we wish to be able to judge the statistical significance of the difference in the means of two independent samples, but where they are simply presented graphically as two (say 95%) Confidence intervals which overlap slightly. from these cam we conclude that the difference is significant at the specified (e.g. 0.05) level?

a what is the probability the 2 observations from a $N(\mu, \sigma)$ distribution will be more than 1.96+1.6 = 3.92 σ's apart? hint; This has the same structure as Q5.38 in M&M edition 2.

b What is the probability that 2 sample means, each having a $N(\mu, \sigma/\sqrt{n})$ sampling distribution, will be more than $3.92(\sigma/\sqrt{n})$'s apart?

-14- Are all head sizes alike?

The following table, from a 1978 article by Epstein, is discussed by Stephen Jay Gould in his book "The Mismeasure of Man". Gould read the original article and found that "a glance at Hooton's original table reveals that the wrong column (standard errors of the mean) had been copied and called standard deviation".

Table 3.2 Mean and standard deviation of head circumference for people of varied vocational statuses

Vocational Status	N	Mean (in mm)	S.D.
Professional	25	569.9	1.9
Semiprofessional	61	566.5	1.5
Clerical	107	566.2	1.1
Trades	194	565.7	0.8
Public service	25	564.1	2.5
Skilled trades	351	562.9	0.6
Personal services	262	562.7	0.7
Laborers	647	560.7	0.3

Source: E A. Hooton, The American Criminal, vol. 1 (Cambridge, Mass.: Harvard Univ. Press, 1939), Table VIII-17.

a Correct the entries in the S.D." column so that the "S.D." column heading is accurate. [If you want to avoid calculations, just change the column heading itself].

b If you were not told about the mislabeling, what features of the data in the table would have led you to suspect that something was wrong.

-15- Sex and sample size

Comment on the statistical argument in the following letter to the Lancet

> A preliminary analysis of data obtained during Britain's first large national survey of sexual attitudes and lifestyles has lately been published.[1] This survey was designed to estimate the size of the HIV epidemic and to collect data that could be used to formulate effective preventive strategies. The results suggest that certain groups of the general population report high numbers of sexual partners. It remains possible, therefore, that a heterosexually transmitted HIV epidemic could occur in the UK. Longitudinal surveys to assess temporal trends in sexual behaviour are essential, and ideally such surveys should gather information about HIV status and risk behaviour. To monitor heterosexual transmission and to measure seroconversion rates, it may be more effective to follow cohorts of young single individuals who report high numbers of sexual partners.[2] However, it is also necessary to monitor risk behaviour changes in the general population. It would obviously be extremely costly to repeat a large-scale national sex survey (this one had a sample size of 18 876) every few years, so can risk behaviour changes be monitored with a much smaller sample size?
>
> I compared some results of this national survey with results of a 1987 pilot study for the national survey. The sample size of the pilot study was 780 individuals, and the results have been analyzed and discussed in detail elsewhere.[2,3] I compared mean values for two sexual behavioural variables (the reported number of sexual partners in the past 5 years, and in the past year) for the two studies, because the frequency distributions of these risk behaviours are very

skewed and the means are very sensitive to skewness (hence, if the frequency distributions are very different, then the means would also be expected to be considerably different).[2,3] Mean values for the two variables are surprisingly similar (table), even though the sample size of the pilot study was only about 1 in 25 of the sample size of the full survey. **These results suggest that it may be possible to monitor temporal changes in certain risk behaviours in the general population by obtaining data from a fairly small sample of individuals every few years.**

SEXUAL PARTNERS IN SURVEYS

	Men		Women	
	National	Pilot	National	Pilot
Mean # partners in past 5 years	2.6	2.9	1.5	1.3
Mean # partners in past year	1.2	1.1	1.0	0.9
Sample size	8384	337	10 492	443

SALLY BLOWER
Department of Biomedical and Environmental Health Sciences, University of California Berkeley, CA 94720

1. Johnson AM *Nature* 1992; 360:410-12. 2. Blower SM. *Philos Trans R Soc Lond (Biol)* 199; 339:33-51. 3. Blower SM*J Acquir Immune Defic Syndr* 1990; 3:763-72.

The Lancet Vol 341 April 17, 1993 page 1023

-16- Two steps forward, one step back [prev exam]

A snail (Schnecke/escargot) starts out to climb a wall. During the day it moves upwards an average of 22 cm (SD 4 cm); during the night, independently of how well it does during the day, it slips back down an average of 12 cm (SD 3 cm). The forward and backward movements on one day/night are also independent of those on another day/night.

a After 16 days and 16 nights, how much vertical progress will it have made? Answer in terms of a mean and SD. Note that -- contrary to what many students calculate -- the SD of the total progress made is not 80 cm; show that it is in fact 20 cm.

b What is the chance that, after 16 days and 16 nights, it will have progressed at least 150 cm?

c Over and above the assumption of independence, which was 'given', did you have to make strong [and possibly unwarranted] distributional assumptions in order to answer part b? Explain carefully, giving justification.

-17- A Close Look at Therapeutic Touch

[Rosa L et al., *JAMA*. 1998;279:1005-1010;
for those interested, there is considerable follow-up correspondence]

See the full article under Other Resources for Chapter 5.

In the last paragraph of Methods the authors state (underlining by JH):

> *The odds of getting 8 of 10 trials correct by chance alone is 45 of 1024 (P=.04), a level considered significant in many clinical trials. We decided in advance that an individual would "pass" by making 8 or more correct selections and that those passing the test would be retested, although the retest results would not be included in the group analysis.*

a Use statistical software, or Table C of M&M3, or first principles, to verify that the probability of getting exactly 8 of 10 correct is indeed 45 of 1024.

b In the next sentence the authors state that in fact they used "8 or more correct" as their criterion. Explain why this definition of "evidence for the therapeutic touch" (or, if you prefer, "against the skeptic's null hypothesis") is more logical than the exactly 8" for which they calculate the P=0.04.

[*Hint*: See the second half of the first paragraph (about *specific* outcomes) under P-values in M&M page 457. In our context, imagine that there were 400 trials: then the probability of -- by chance alone -- getting exactly 320 is indeed, in Dr. Arbuthnot's words, "vanishingly small". but the probability of getting specifically 200 (a value that provides no evidence against H_0, is also small (0.04)]

c Calculate -- under the "null" hypothesis, the probability of "8 or more correct". Is it indeed less than the arbitrary "level considered significant" of 0.05? If not, then what would the criterion need to be so that the probability -- again calculated under "H_0" -- of reaching this criterion is < 0.05.

d Figure 2 shows the scores of the 28 subjects. Multiply the set of Binomial probabilities with n=10 and $\pi = 0.5$ (i.e., p[0/10 correct | $\pi = 0.5$] to p[10/10 correct | $\pi = 0.5$] by 28 to obtain theoretical frequencies. These are the numbers of subjects, out of 28, one would expect to get 0/10, 1/10, ... 10/10 trials correct if all they were doing in each trial was guessing. Compare the theoretical frequencies of subjects with the observed "No. of subjects" with each score. Comment.

Ignore for the moment the fact that the 28 people tested were really only 21 distinct people -- 14 tested once (10 trials each) and 7 tested twice (10 trials, twice)