

PAD 504 - Final Exam

Please answer each of the following in the bluebooks provided. Show all significant work for part credit. Do not tear out pages from your bluebook — just draw a line through work you want me to ignore.

(1) [8 points] Short answer quickies:

- (a) [2] A witness to a crime tries to identify the three people involved. Suppose she has a 90 percent chance of being right on each independent identification. What is the probability that she is wrong on at least one of the three?
- (b) [2] In a multiattribute decision, what are "range sensitivity" and "sensitivity to reasonable changes in weights" and why are each important?
- (c) [2] In the MAU model begun at the right, the modeler wants to put *formulas* in row 7 that can be "filled down" to rows 8 and 9 to scale the data in rows 2 through 4 on linear scales from 0 to 100. What cell formula would you put in cell C7?
- (d) [2] If you roll one die and draw one card from a deck, what is the probability that you will roll a 1 and draw an ace? What is the probability you will roll a 1 or draw an ace? (Here, as usual, "or" means one or the other or both.)

(2) [12 points] You have been asked to create a dynamic model in a spreadsheet capturing enrollments in the MPA program in the Rockefeller College.

You find that in 1994 there were 80 students in their first year of the MPA, 76 in their second year, 27 in their third year, and 10 who were in their fourth year or beyond (pursuing the degree part-time). During that year 80 students were admitted to begin the program in 1995. From previous years you determine that 5 percent of the students tend to drop out by the end of their first year; 65 percent of the second year students graduate or drop out after their second year (meaning that 35 percent remain); 70 percent of the third year students graduate or drop out; and 80 percent of those in the program for four or more years graduate or drop out. The department plans on admitting 80 students each year to begin the program the next year.

(a) [8] Build a spreadsheet model in your bluebook capturing this information. Please formulate your model as shown below (and please make your spreadsheet large enough so that it is very easy for me to read). Show *numbers* in column B and *cell formulas* in columns C and D.

Year	1994	1995	1996
First year students			
Second year students			
Third year students			
Four or more years			
Total			
Entrants for next year			

(b) [4] Describe how you would use your model to determine how large the total enrollment in the MPA would grow if the department began admitting 100 students in 1996 to begin the program in 1997. [Don't bother with any computations!]

(3) [12 points] Before the New York State mandatory seatbelt law was adopted a group of 172 people were asked whether they agreed with the proposed law or were opposed to it. The table at the right shows the numbers of people in the various opinion categories, tabulated with their political parties.

(a) [6] Compute the number of observations in each cell that would be expected if opinion on the seatbelt law were independent of political party.

(b) [3] Compare the observed frequencies with your expected values. Would you say there is some association in this data between opinion on the seatbelt question and political party, or would you say the two are largely independent? Defend your answer.

(c) [3] Show how you could build a spreadsheet that would compute all the expected values, starting as follows:

	Democrat	Republican	Totals:
Str agree			67
Agree			52
Neutral			17
Disagree			18
Str disagree			18
Totals:	126	46	

(4) [10 points] The box-and-whiskers plots at the right show the number of marriages and divorces in 1979 in 47 of the 48 continental United States.

(a) [6] Name, and give a value for, all the numbers can you read from the divorce plot.

(b) [2] What can you say from the marriage plot about the shape of distribution of the marriage data?

(c) [2] What can you say about the mean of the marriage data? What can you say about the mean of the divorce data?

(5) [12 points] The Foundation you work for funds its projects from the annual revenue from its investments. You have \$100,000 to invest in stocks of three possible companies: ATT, BTT, and CTT, and you want to determine how many shares of each stock the Foundation should buy.

The projected annual revenue per share is \$7 for ATT, \$3 for BTT and \$3 for CTT. The cost per share is \$60 for ATT, \$25 for BTT, and \$20 for CTT. To be sure your spread the investment, you want to invest no more than \$60,000 in ATT, \$25,000 in BTT and \$30,000 in CTT. The question is: How many shares of each stock should your Foundation buy to maximize its annual revenue from these stocks? You won't be able to answer that question on this exam, but you can show how to set up the problem to solve it as an exploratory optimization (linear programming) problem.

Copy the following spreadsheet for this problem in your bluebook (nice and large) and fill in all the appropriate cells with data and cell formulas. (You need not, and should not, do any calculations.)

	<u>AT&I</u>	<u>BT&I</u>	<u>CT&I</u>
Shares:			
Revenue/share:			
Cost/share:			
Total revenue =	?		
Maximize annual revenue subject to:			
	Amount	Maximum	Slack
Cost constraint:			
Maximum in ATT:			
Maximum in BTT:			
Maximum in CTT:			

(6) [10 points] Child care experts in Stokey City believe that about 3 percent of the city's 20,000 children are physically abused. They propose a screening program designed to identify abused children so they and those around them can be helped. The screening process is not perfect: if a child is actually abused, the probability is 0.95 that the screening test will be positive (that is, indicate abuse); if the child is not abused, the probability that the test will be incorrectly positive is 0.10.

The probability tree on the left captures this information.

School officials want to find the probability that a child is abused or not, given that the screening test shows a positive or negative result. Which is to say, they need the tree flipped around, like the one on the right above.

Find the probabilities the school officials need by answering the following:

- (a) [2] Find in tree A the four branch probabilities:
 - P(Abused and Test positive)
 - P(Abused and Test negative)
 - P(Not abused and Test positive)
 - P(Not abused and Test negative)

(b) [2] Transfer your results from (a) to find the four branch probabilities in tree B:

- P(Test positive and Abused)
- P(Test positive and Not abused)
- P(Test negative and Abused)
- P(Test negative and Not abused)

(c) [2] From (b), compute the probabilities of children in Stokey City testing positive and negative, that is compute $P(\text{Test positive})$ and $P(\text{Test negative})$ in tree B.

(d) [4] Use (b) and (c) to compute the remaining probabilities in tree B:¹

- $P(\text{Abused} \mid \text{Test positive})$
- $P(\text{Not abused} \mid \text{Test positive})$
- $P(\text{Abused} \mid \text{Test negative})$
- $P(\text{Not abused} \mid \text{Test negative})$

[It turns out here that *false* positives are more likely than *true* positives, so if you get such an unhappy result don't assume you've made an error somewhere -- you may just be right.]

(7) [14 points] Estimates vary about the costs of national health care reform, and experts have different levels of confidence in their estimates.

(a) [4] Compute the expected costs for each of the plans in the decision tree at the right. Which plan would an EMV'er pick?

(b) [6] Senator Rohrbaugh objects to a decision tree here because he says judging a health care plan is really a multiattribute decision, not a judgment to be made on the one criterion of costs, even if they are adjusted by good probability estimates.

How might you set up a multiattribute utility model to aid in deciding among the competing national health care proposals? Be as specific as your knowledge of MAU models and national health care policy criteria allow. (If you can't think of real criteria, make up some plausible ones — this is a question about MAU models, not health care.)

(c) [4] Discuss Senator Rohrbaugh's claim. Do you agree with him? Explain the pros and cons.

¹ Remember that $P(\text{Abused} \mid \text{Test positive})$ means the "probability that the child is Abused given that the child tested positive."

Answers

year	1994	-13561	-13305
First yr	80	20480	20480
Second yr	76	19456	19456
Third yr	27	-1.#INF	-1.#INF
Fourth or more	10	1.#INF	#N/A!
Total	-16128	#N/A!	1.#INF
Entrants	80	20480	20480

	Democrat	Republican	Totals:
Str agree	-1.\$	-1.\$	67
Agree	-1.\$	#N/A!	52
Neutral	#N/A!	1.\$	17
Disagree	1.\$	#N/A!	18
Str disagree	1.\$	#N/A!	18
Totals:	126	46	-21504

	ATT	BTT	CTT
Shares	910	1000	1000
Revenue/share	\$7	\$3	\$3
Cost/share	\$60	\$25	\$20
Total revenue	\$21040		
Maximize annual revenue subject to:			
	Amount	Maximum	Slack
Cost constraint:	\$277152000	\$100000	-\$28671
Max in ATT	\$1221918720	\$60000	\$6165
Max in BTT	-\$22431	\$25000	\$0
Max in CTT	\$8270	\$30000	\$4135

Observed Frequency Table

Clinton plan

Data	<u>Cost</u>	<u>Benefit</u>
Option X	12	33
Option Y	24	48
Option Z	28	60
Evaluations	<u>Cost</u>	<u>Benefit</u>
Option X		
Option Y		
Option Z		