

This is an open book exam. You are allowed to use your text, handouts and notes. Other books, laptops, PDA's and text messaging devices are prohibited. Calculators are permitted. Be sure to give complete explanations to receive full credit. There are [150] total points. **Do SEVEN of nine problems.** If you do more than seven, the first seven will be graded. Cross out the problems you don't wish to be graded.

1.	_____	/21
2.	_____	/21
3.	_____	/21
4.	_____	/21
5.	_____	/21
6.	_____	/21
7.	_____	/22
8.	_____	/22
9.	_____	/22
Total		_____/150

1. [21] At Milford Mills, three machines  $M_1$ ,  $M_2$  and  $M_3$  make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 1%, 2%, and 3% of the products made by each machine, respectively, are defective. If a finished product were chosen randomly and found to be defective, what is the probability that it is made by machine  $M_3$ ?

Your grade will be posted at my office according to

*Secret I.D. :*

---

2. In a study conducted by Wright State University, senior high school students were asked if they ever used marijuana. Of 1120 male students, 445 said yes. Of 1156 female students, 515 said yes. Assuming that the observations can be treated as a random sample from a population of interest, determine whether there is a significant difference in the proportion of male and female students who said they have used marijuana.

(a) [6] State the null and alternative hypotheses. State the formula for the test statistic. What assumptions on the data are required for this to be a valid statistic?

(b) [15] Compute the  $P$ -value. What conclusions can you draw at the 0.05 level?

3. [21] Suppose  $X_1, X_2, \dots, X_{100} \sim \text{Exp}(\lambda = 1)$  is a random sample taken from an exponential distribution with parameter  $\lambda = 1$  and  $Y_1, Y_2, \dots, Y_{200} \sim U(0, 1)$  is a random sample taken from a uniform distribution on the interval  $[0, 1]$ . If all  $X_i$ 's and  $Y_j$ 's are mutually independent, find the probability

$$P(S_X + S_Y < 212)$$

where  $S_X = X_1 + \dots + X_{100}$  and  $S_Y = Y_1 + \dots + Y_{200}$ .

4. In a study of environmental poisons (“The Accumulation and Excretion of Heavy Metals in Organisms,” *Heavy Metals in the Aquatic Environment*, P. Krenkel, ed., Oxford Press, pp. 155-162, 1975) volunteers ingested methylmercury to test whether there is a difference in hazard to men and women. Test at the 0.01 level of significance whether there is strong evidence that the average half-lives differ in men and women.

	Methylmercury Half-Lives (in Days)										n	Mean	S. D.
Females	52	69	73	88	87	56					6	70.83333	15.09194
Males	72	88	87	74	78	70	78	93	74		9	79.33333	8.077747

- (a) [8] State the null and alternate hypotheses. State the formula for the test statistic. Under what assumptions is your test statistic valid?

- (b) [13] Perform the calculations. What is your conclusion?

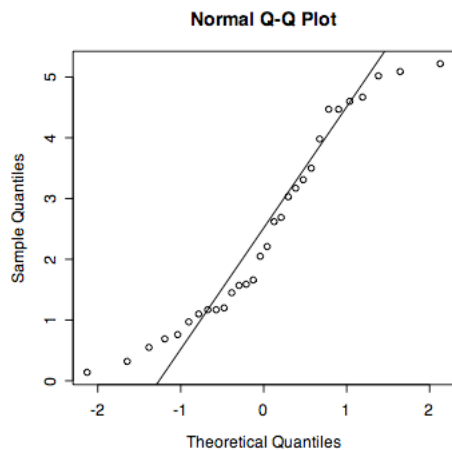
5. In a study (“Engineering Properties of Soil,” *Soil Science*, pp. 93–102, 1998) a random sample of soil specimens was selected and the amount of organic matter was determined for each specimen. Does this data strongly suggest that the average amount of organic matter is other than 3.00%? The data was analyzed in R ©.

```
R : Copyright 2005, The R Foundation for Statistical Computing  
Version 2.1.1 (2005-06-20), ISBN 3-900051-07-0
```

```
> soildata <- c(1.10, 5.09, 0.97, 1.59, 4.60, 0.32, 0.55, 1.45,  
+ 0.14, 4.47, 1.20, 3.50, 5.02, 4.67, 5.22, 2.69,  
+ 3.98, 3.17, 3.03, 2.21, 0.69, 4.47, 3.31, 1.17,  
+ 0.76, 1.17, 1.57, 2.62, 1.66, 2.05)  
> qqnorm(soildata)  
> qqline(soildata)  
> t.test(soildata, mu=3.00)
```

One Sample t-test

```
data: soildata  
t = -1.7583, df = 29, p-value = 0.08924  
alternative hypothesis: true mean is not equal to 3  
95 percent confidence interval:  
 1.878043 3.084623  
sample estimates:  
mean of x  
 2.481333  
  
> sd(soildata)  
[1] 1.615641
```



- (a) [5] What assumptions on the data must hold to do the t-test? Do the assumptions hold for this data? Explain.
- (b) [8] State the null and alternative hypotheses. What conclusions do you draw at the 0.05 level of significance?
- (c) [8] Suppose that the population mean was in fact 2.00 and the population standard deviation was 1.616. What is the probability at the 0.05 level of concluding that the evidence is not strong that the mean differs from 3.00%? How large a sample is needed so that this probability is at most 0.05?

6. A recent study (*J. Amer. Med. Assoc.*, vol. 284, p. 831, 2000) considered whether daily consumption of 1200 mg. of garlic could reduce tick bites. The treatment was administered to 20 Swedish military conscripts. Suppose that the garlic did better than placebo in reducing tick bites for 12 of the subjects.

(a) [13] Find a 90% lower confidence bound for the proportion of people for whom the garlic treatment did better than placebo in reducing tick bites.

(b) [8] Is there significant evidence that the garlic treatment does better than placebo more than 50% of the time? State the null and alternate hypotheses. Using only your results from part (a.), what is your conclusion?

7. A manufacturer of salad dressings needs to determine whether the new filling machine is dispensing the proper amount, 8 ounces per bottle. Periodically, a sample of 50 bottles is randomly selected, and the filling line is stopped if there is strong evidence that the mean amount dispensed is different than 8 ounces. Is there strong evidence that the average is different than 8 ounces? Suppose that the average amount dispensed in a particular sample of 50 bottles was 7.983 ounces with a standard deviation of 0.150 ounces.

(a) [5] State the null and alternative hypotheses. State the formula for the test statistic and the acceptance and rejection regions. What assumptions on the data are required for your statistic to be appropriate?

(b) [8] Using a 0.05 level of significance, is there strong evidence that the average is different than 8 ounces?

(c) [7] What is the probability that the test in part (a.) will show that the average fill is not significantly different from 8 ounces, when in fact the true population mean is 7.952 ounces and population standard deviation is 0.150 ounces?

8. In a study of 13 children suffering from asthma (*Clinical and Experimental Allergy*, vol. 20, pp. 492–432, 1990), single doses of formoterol (F) and salbutamol (S) were compared. The outcome measured the child's peak expiratory flow (PEF) eight hours after the treatment. Does the data strongly indicate that there is a difference in the PEF for the two treatments? The data on PEF follow:

Child	F	S	Child	F	S	Child	F	S	Child	F	S
1	310	270	5	410	380	9	330	365	13	220	90
2	385	370	6	370	300	10	250	210			
3	400	310	7	410	390	11	380	350			
4	310	260	8	320	290	12	340	260			

---

- (a) [6] State the null and alternative hypotheses. Give the formula for the test statistic. What assumptions are you making about the data in order that your statistic be appropriate?

- (b) [17] Compute the  $P$ -value and state your conclusions.



9. [23] An insurance company offers discounts to homeowners who install smoke detectors in their homes. A company representative claims that 70% or more of the policyholders have smoke detectors. Suppose that you wish to design an experiment to test the rep.'s claim. Of over 5000 policy holders, you will randomly select a certain number  $n$  of the policyholders and check whether they have smoke detectors. If at most  $m$  of them don't own smoke detectors, then you will reject the claim that 70% have smoke detectors. Approximately, what is the smallest  $n$  and what is the corresponding  $m$  so that the probability of rejecting the claim is at most 0.20 when the actual proportion of households with smoke detectors is 70%, and that the probability of accepting the claim is at most 0.20 when in fact the actual proportion with smoke detectors is 50%?

