1. (10%) A new test has been devised for detecting a particular type of cancer. If the test is applied to a person who has this type of cancer, the probability that the person will have a positive reaction is 0.95 and the probability that the person will have a negative reaction is 0.05. If the test is applied to a person who does not have this type of cancer, the probability that this person will have a positive reaction is 0.05 and the probability that the person will have a negative reaction is 0.95. Suppose that in the general population, one person out of every 100,000 people has this type of cancer. If a person selected at random has a positive reaction to the test, what is the probability that he has this type of cancer?

2. (10%) Suppose Y is a random variable with mean = μ and variance = σ^2. Let's define a new random variable A and $A_k = \lambda Y_k + (1-\lambda)A_{k-1} \quad (k = 1, 2, \cdots k)$ ($0 < \lambda < 1$). It can be proved that A_k is an exponentially weighted moving average of all Y values and, $A_k = \sum_{i=0}^{k} \lambda(1-\lambda)^i Y_{k-i}$
Variance of A = ? (assume k is very large)
(Hint: $1 + w + w^2 + w^3 + \ldots = 1/(1-w)$ when $0 < w < 1$)

3. (10%) Two methods are used to predict the shear strength for steel plate girders. Data for nine girders in the form of the ratio of predicted to observed load for these two methods are as follows:

Girder	Method 1	Method 2	Difference
1	1.186	1.061	0.125
2	1.151	0.992	0.159
3	1.322	1.063	0.259
4	1.339	1.062	0.277
5	1.200	1.065	0.135
6	1.402	1.178	0.224
7	1.365	1.037	0.328
8	1.537	1.086	0.451
9	1.559	1.052	0.507
	$\overline{Y} = 1.3401$ $S = 0.1460$	$\overline{Y} = 1.0662$ $S = 0.0494$	$\overline{d} = 0.2739$ $S_d = 0.1351$

Construct a 95 percent confidence interval for the difference in mean predicted to observed load.

4. (20%) An experimental design was used to study the possible effects of three factors on the output response.

Test	Factor 1 (x_1)	Factor 2 (x_2)	Factor 3 (x_3)	Response Y	
				Trial 1	Trial 2
1	-1	-1	-1	9	11
2	1	-1	-1	18	22
3	-1	1	-1	4	4
4	1	1	-1	9	11
5	-1	-1	1	7	9
6	1	-1	1	17	19
7	-1	1	1	5	7
8	1	1	1	10	14

Assume there is no interaction effect ($Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \varepsilon$), you can construct an ANOVA table to identify which of the three factors are significant.

(a) (10%) Sum of squares due to factor 1 = ?

(b) (10%) Degrees of freedom for error = ?

5. (20%) Suppose that two independent measurements X and Y are made of the rainfall during a given period of time at a certain location and that the p.d.f. g of each measurement is as follows:

$$g(x) = \begin{cases} 2x & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases},$$

(a) Determine the value of P(X + Y ≤ 1)

(b) Let Y = $X^{1/2}$, then E(Y) = ?

6. (10%) Suppose that X_1, X_2, and X_3 are independent random variables such that $E(X_i) = 0$ and $E(X_i^2) = 1$ for i= 1, 2, 3. Determine the value of $E[X_1^2(X_2 - 4X_3)^2]$

7. (10 points) A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hour of travel. The second door leads to a tunnel that will return him to the mine after 5 hours of travel. The third door leads to a tunnel that will return him to the mine after 7 hours of travel. If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

8. (10 points) The computer output for the multiple regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

is shown below. However, because of a printer malfunction some of the results are not shown. These are indicated by the boldface letters ***a*** to ***j***. Fill in the missing results (up to three decimal places).

Predictor	*Coef*	*StDev*	*T*
Constant	25.277	***a***	4.11
x_1	3.51	2.808	***b***
x_2	***c***	0.30	-2.367

S = ***d*** R-Sq = ***e***

ANALYSIS OF VARIANCE

Source of Variation	*df*	*SS*	*MS*	*F*
Regression	***g***	412	***i***	***j***
Error	***h***	***f***		
Total	39	974		

試題請隨卷繳回