**1.** For a disease known to have a postoperative complication frequency of 20%, a surgeon suggests a new procedure. He tests this new procedure on 10 patients and found that there are no complications. What is the probability that in operating on 10 patients with the traditional method

(i) no patient will suffer from complication

(ii) less than 5 patients will suffer from complication

(iii) more than 5 patients will suffer from complication

2. Assume that lifetimes of a certain medical device follow a Gaussian distribution with a mean of 300 hours and standard deviation of 25 hours.

(i) Find the probability that any one of these medical devices have a lifetime of more than 325 hours.

(ii) What percentage will have lifetimes of 300 hours or less?

(iii) What percentage will have lifetime between 220 and 260 hours?

3. In clinical trial, a newly designed drug was given to ten patients, and the increments in their blood pressure were recorded to be 2, 6, -2, 4, -3, 6, 4, 0, 0, 3. Is it reasonable to believe that the drug has no effect on blood pressure?

4. Suppose that a researcher measured the biomass (in milligrams) produced by bacterium A and bacterium B, in shake flasks containing glucose as substrate. The researcher had 4 replicate flasks of each bacterium as given below:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Replicate 1** | | **Replicate 2** | | **Replicate 3** | | **Replicate 4** | |
| **Bacterium A (mg)** | 520 | | 460 | | 500 | | 470 |
| **Bacterium B (mg)** | 230 | | 270 | | 250 | | 280 |

Which bacterium produces more biomass when grown on glucose?

5. Given below are the weight gains of baby chicks (grams) under 4 different feed treatments (t1, t2, t3 and t4).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Treatments | | | |
| Observation | T1 | T2 | T3 | T4 |
| 1 | 55 | 61 | 42 | 169 |
| 2 | 49 | 112 | 97 | 137 |
| 3 | 42 | 30 | 91 | 169 |
| 4 | 21 | 89 | 95 | 85 |
| 5 | 52 | 63 | 92 | 154 |

Do the four treatments differ significantly in respect of weight gain?

**6.**  The ages (in years) and blood pressures of 9 men are given below:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Age (X):** | 36 | 47 | 49 | 42 | 60 | 72 | 55 | 41 | 63 |
| **BP (Y):** | 117 | 127 | 144 | 138 | 154 | 157 | 142 | 124 | 148 |

(i) Find the correlation coefficient between X and Y and test for its significance

(ii) Find the regression equation of Y and X

(iii) Estimate the blood pressure of a man of 40 years

7. A survival data contain remission times (in weeks) of 42 multiple leukemia patients in a clinical trial, half of whom get a certain new treatment therapy and the other half of whom get a standard treatment therapy. The exposure variable of interest is treatment status (*Rx* = 0 if new treatment, *Rx* = 1 if standard treatment). Two other variables for control as potential confounders are log white blood cell count (i.e., *logwbc*) and *sex*. Failure status is defined by the variable *Relapse* (0 if censored, 1 if failure). This data set is available for download here.

(i) Suppose we wish to describe Kaplan-Meier (KM) curves for the variable *logwbc*. Because *logwbc* is continuous, we need to categorize this variable before we compute KM curves. Suppose we categorize *logwbc* into three categories - low, medium, and high – as follows:

low (0–2.30), *n*= 11;

medium (2.31–3.00), *n* = 14;

high (>3.00), n = 17.

Based on this categorization, plot KM survival curves for each of the three categories of *logwbc*. Also, comment on how the three curves compare with each other.

(ii) Perform the log–rank test for three categories created above in (i), and draw conclusions from the test.

(iii) Perform a univariate Cox regression with explanatory variable *as Rx*, *sex* or *logwbc*. Now, perform a multivariate Cox regression including all three variables and interpret the result.