1. Consider Multinomial distribution with $p_1 = \frac{1}{8}$, $p_2 = \frac{1}{2}$ and $p_3 = \frac{3}{8}$
   (a) Simulate 100 samples from Multinomial with size 30 and compute the Pearson-$\chi^2$ statistic for each repetition, saving it in a variable `Xsquared`.
   (b) On the same graph: plot the histogram of `Xsquared` and chi-squared density with 3 degrees of freedom.

2. The students hostel october mela has a game involves rolling 3 dice. The winnings are directly proportional to the total number of ones rolled. Suppose Jooa brings her own set of dice and plays the game 100 times. Her results are tabulated below:

| Number of ones | Number of Rolls |
|---|---|
| 0 | 40 |
| 1 | 37 |
| 2 | 13 |
| 3 | 10 |

   Student's supreme leader, Moola ,gets suspicious. He wishes to determine if the dice are fair or not. Let us help him with the $\chi^2$-square goodness of fit test.
   (a) Compute the respective Multinomial probabilities in R.
   (b) Plot the barplot of the observed counts and expected counts.
   (c) Plot the barplot of the $\frac{\text{observed count- expected count}}{\sqrt{\text{expected count}}}$
   (d) Compute the chi-square statistic and decide if the null hypothesis that the dice are fair can be rejected or not ?

3. Consider the data from your 50 rolls of dice in August 6th Class Worksheet
   (a) Create a data vector called `rolls` and create a vector of length 6 called `counts` that provides number of times $1, 2, 3, 4, 5, 6$ occured in the 50 rolls.
   (b) Using $\chi^2$-goodness of fit test can you decide if the die you rolled was fair ?

1. Take the data `dengueb.csv` from the shared folder.
   (a) Categorise the data into three groups assigning Marker to be $0, 1, 2$ depending on the values of `BICARB` being less than or equal to 16, between 16 and 21, and greater than 21.
   (b) Let the Null hypothesis be that the variables are independent. Compute the `Xsquared` statistic for the data and the p-value (i.e Probability that chi-squared random variable with appropriate degrees of freedom exceeds the observed `Xsquared`) for the null hypothesis.
2. Take the dataset `heartrate` from the package `UsingR`.
   (a) Make a scatter plot of the variables `maxrate` and `age`. Using the inbuilt `lm` function plot the regression line of `maxrate` $= \beta_0 + \beta_1$`age` on the same plot.
   (b) Using the inbuilt `resid` function compute the residuals at each point. Using the inbuilt `fitted` function compute the fitted values at each point. Verify that `resid` = actual values of data - `fitted`.
   (c) Using the generic `predict` function or otherwise, from the above line find the `maxrate` for `age` $= (12, 33, 45, 76)$.
   (d) Write a simple `function` to test the null hypothesis that $\beta_1 = -1$ versus the alternative that $\beta_1 \neq -1$. Decide if the data has evidence to reject the null hypothesis.
   (e) From the estimate of $\hat{\beta}_1$ obtained above find a 95% confidence interval for $\beta_1$.
   (f) Do the following plots:
      i. Scatter plot of residuals versus fitted values.
      ii. Scatter plot of $\sqrt{\text{standardised residuals}}$ versus fitted values.
      iii. A NormaaQ-Q plot for the residuals.
3. Repeat question 2(a), 2(b), and 2(f) for the dataset `phones` from the `MASS` package. In