1. A random sample of $m = 10$ women and $n = 20$ men is gathered, and each person is tested for a disease. We assume for this problem that the test is completely accurate. The numbers of women and men in the sample who have the disease are X and Y respectively, with $X \sim \text{Binomial}(m = 10, p = 0.1)$ and $Y \sim \text{Binomial}(n = 20, p = 0.1)$, independently.

   Answer the following questions:

   (a) Calculate $\mathbb{P}(X = 1, Y = 2)$. Quote your numerical answer to 4 decimal places.

   (b) Calculate $\mathbb{P}(X + Y = 3)$. Quote your numerical answer to 4 decimal places.

   (c) Let $R = X + Y$. Note that R represents the number of the people in the sample that have the disease. Find the conditional pmf of X given that $R = 3$.

   (d) Use the result that you have obtained in part (c) in order to compute $\mathbb{E}(X|R = 3)$.

   (e) Compute $\mathbb{E}(X|R = 3)$ under the hypothesis that $p = 0.2$ instead of $p = 0.1$.

2. I do three different experiments:

   - Experiment$_1$: I toss a coin 104 times and get 81 heads.
   - Experiment$_2$: I toss a coin 84 times and get 68 heads.
   - Experiment$_3$: I toss a coin 86 times and get 57 heads.

   Answer the following questions:

   (a) Let X_1, X_2 and X_3 be the number of heads in Experiment$_1$, Experiment$_2$ and Experiment$_3$, respectively. In all experiments, the null hypothesis is that the coin is fair. Under the null hypothesis, state the distributions of X_1, X_2 and X_3.

   (b) In each experiment, we wish to test the hypothesis that the coin is fair. For Experiment$_1$, formulate the null hypothesis and alternative hypothesis, in terms of the distribution of X_1 and its parameters. Remember to specify the full distribution of X, and use a two-sided alternative hypothesis.

   (c) For each experiment, write down the `R` command required to find the p-value for the hypothesis test, and run this command in `R` to find the p-value.

   (d) Based on the p-values computed in part (c), decide the experiment in which we have the weakest evidence against the null hypothesis.

3. Let $X \sim \text{Geometric}\left(\dfrac{1}{\lambda+1}\right)$, where $0 < \lambda < \infty$.

   Answer the following questions:

   (a) For the particular case when $\lambda = 1$, calculate $F_X(1.99)$ and $\mathbb{P}(X > 2)$.

   (b) In the general case when $\lambda \in (0, \infty)$, show that

   $$\mathbb{P}(X = k + t | X \geq k) = \mathbb{P}(X = t), \text{ for all nonnegative integers } k \text{ and } t.$$

   For ease of computation, you may use the notation $p = \dfrac{1}{\lambda + 1}$. It is obvious that $0 < p < 1$.

(c) Let $x_1, \ldots, x_n$ be *independent* observations for the random variable X. Suppose that $n > 1$ and $x_1 + x_2 + \cdots + x_n > 0$. Assume that the parameter λ is unknown. Show that the log-likelihood function has the expression:

$$\log L(\lambda; x_1, \ldots, x_n) = -n(\bar{x} + 1)\log(\lambda + 1) + n\bar{x}\log(\lambda),$$

where $\log(\cdot) = \log_e(\cdot)$ denotes the natural logarithm and $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n}$.

(d) Use the result in part (c) to obtain the maximum likelihood estimate $\hat{\lambda} = \bar{x}$. In your answer, you can refer to Figure 1.

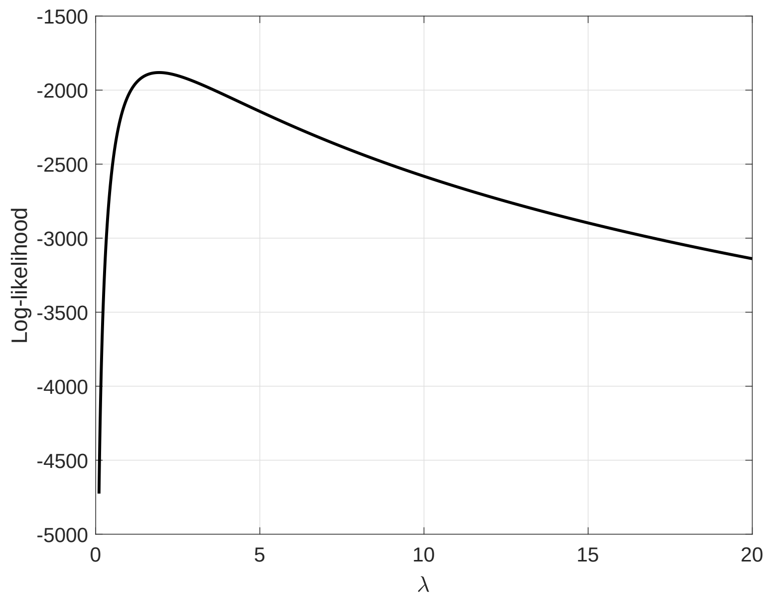

Figure 1: Plot for question 3, part (d). Note that $\bar{x} = 1.93$.