1. **Packets Over the Internet**

   \( n \) packets are sent over the Internet (\( n \) even). Consider the following probability models for the packet loss process:

   (a) Each packet is routed over a different path and is lost independently with probability \( p \).

   (b) All \( n \) packets are routed along the same path, and with probability \( p \), one of the links along the path fails and all \( n \) packets are lost. Otherwise all packets are received.

   (c) The \( n \) packets are divided into 2 groups of \( n/2 \) packets, and each group is routed along a different path and lost with probability \( p \). Losses of different groups are independent events.

In each of the three models, compute the distribution of the number of packet losses. For \( n = 6 \) and \( p = 0.3 \), plot the distribution in each of the three cases. Does the distribution depend on the probability model? Which of the three routing protocols do you prefer?

2. **Coupon Collection**

   Consider the coupon collection problem with \( n \) distinct baseball cards and \( m \) cereal boxes bought. Compute the distribution of the number of Babe Ruth cards acquired.

3. **Family Planning**

   Mr and Mrs Brown decide to continue having children until they either have their first girl or until they have five children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let \( B \) and \( G \) denote the numbers of boys and girls respectively that the Browns have.

   (a) Write down the sample space together with the probability of each outcome.

   (b) Compute and plot the distributions of the random variables \( B \) and \( G \).

4. **Conditional Independence and Medical Diagnosis**

   There is a disease which affects 2% of the population. A medical test is available which gives a false positive rate of 5% and a mis-detection rate of 3%.

   (a) Compute the false discovery rate and the false omission rate of using this test.

   (b) The doctors complain that the false discovery rate for this test is too high. The medical test company responds with a new test that has a false positive rate of 6% and a mis-detection rate of 4%. Although this new test has worse false positive rate and mis-detection rate compared to the old test, the company claims that when used in conjunction with the old test, will give a lower false discovery rate because the results of the two tests are conditionally independent given the disease state of the patient. More specifically, the company recommends the doctors diagnose a patient to have a disease if and only if both the old and the new tests are positive.

   Do you agree that the company’s claim? Justify your answer.

5. **Typos.** Assume a 10-page scientific journal paper has 10 typos in it.

   (a) If the typos are randomly distributed over the 10 pages, find the probability of having a single typo on each page.

   (b) If 8 of the typos are spelling errors and the other two are mathematical errors, find the probability of having 4 spelling errors and a single mathematical error on the first five pages of the paper.