Practice with Binomial random variables

Suppose you roll a die three times. Let \( X \) denote the number of times an even number appears, and \( \hat{p} \) the proportion of times an even number appears.

1. Is \( X \) a binomial random variable? Why or why not? If it is, what are \( n \) and \( p \)?

   Yes. It satisfies:
   1) \( X \) is a count of "success" in a trial
   2) \( n = 3 \) is fixed in advance
   3) each trial is independent of every other
   4) prob. of success constant from trial to trial

   \( p = 1/2 \)

2. What values can \( X \) assume? What values can \( \hat{p} \) assume? Will \( \hat{p} \) ever equal \( p \)?

   \( X \) can be 0, 1, 2, 3

   \( \hat{p} \) can be 0, 1/3, 2/3, 1

   \( \hat{p} \) will not ever equal \( p \).

3. The die is rolled three times, and each time is either a "success" (an even number) or a "failure" (an odd number). The following table lists the full sample space of possible outcomes. For each one, calculate the probability, and the corresponding values of \( X \) and \( \hat{p} \)

   \[ \begin{array}{ccc|cc}
   \text{event} & \text{probability} & X & \hat{p} \\
   \hline
   
   FFP & \frac{1}{8} & 0 & 0 \\
   FFS & \frac{3}{8} & 1 & 1/3 \\
   FSF & \frac{3}{8} & 1 & 1/3 \\
   SFF & \frac{1}{8} & 1 & 1/3 \\
   FSS & \frac{3}{8} & 2 & 2/3 \\
   SFS & \frac{1}{8} & 2 & 2/3 \\
   SSF & \frac{1}{8} & 3 & 3/3 \\
   SSS & \frac{1}{8} & 3 & 3/3 \\
   \end{array} \]

4. Use your answers above to write down the probabilities for each possible value of \( X \) and \( \hat{p} \).

   \[ \begin{array}{cccc}
   x_i & 0 & 1 & 2 & 3 \\
   \text{Prob}(x_i) & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\
   \end{array} \]

   \[ \begin{array}{cccc}
   \hat{p}_i & 0 & 1/3 & 2/3 & 1 \\
   \text{Prob}(\hat{p}_i) & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\
   \end{array} \]

5. Check these probabilities against those in the Binomial table. Are they the same? (If so, you’ve basically just derived where the numbers in the table come from.)

   Yes, they are the same as the probabilities listed in the

   \( n = 3 \), \( p = \frac{1}{2} \) row.
6. Using the data in the tables of problem 4, calculate the mean of \( X \) and of \( \hat{p} \). Are your answers consistent with what you expect from the formulas I gave in class? (If so, you’ve basically derived where these formulas come from.)

\[
\text{mean of } X: \quad 0 \left( \frac{1}{8} \right) + 1 \left( \frac{3}{8} \right) + 2 \left( \frac{3}{8} \right) + 3 \left( \frac{1}{8} \right) = \frac{3 + 6 + 3}{8} = \frac{12}{8} = \frac{3}{2}
\]

\[
\text{mean of } \hat{p}: \quad 0 \left( \frac{1}{8} \right) + \frac{1}{3} \left( \frac{3}{8} \right) + \frac{2}{3} \left( \frac{3}{8} \right) + 1 \left( \frac{1}{8} \right) = \frac{1}{2}
\]

Formulas:

\[
\text{mean of } X = np = 3 \cdot \frac{1}{2} = \frac{3}{2} \quad \text{and} \quad p = \frac{1}{2}
\]

\[
\text{Consistency.}
\]

7. Make a probability histogram for \( X \) and \( \hat{p} \). (Recall that a probability histogram is just a bar chart, with the values of the variable on the \( x \)-axis and bars representing probabilities extending up in the \( y \)-direction.) Does the shape of these histograms seem vaguely bell-like?

![Probability Histograms](image)

Yes, bell-like.

8. Now imagine you rolled the die 100 times. Based on your work above, take a guess as to how the probability histogram for \( X \) would look. Draw a rough sketch. (Your sketch can be crude, i.e. you don’t need to draw all 101 bins on the \( x \)-axis.)

![Probability Histogram Sketch](image)
Practice with Sample Counts and Proportions

1. Numerous studies indicate that about 25% of jazz fans prefer vinyl to compact disc. Suppose you take a survey of 20 jazz fans.

   (a) What is the distribution of the number that prefers vinyl?

   \[ X \sim \text{B}(20, 0.25) \]

   (b) What is the probability that at least 6 of the people interviewed prefer vinyl?

   \[
   P(X \geq 6) = 1 - P(X \leq 5) \\
   = 1 - \left[ .0032 + .0211 + .0669 + .1339 + .1897 + .2331 \right] \\
   = .383
   \]

2. Suppose in the problem above you had interviewed 1000 jazz fans.

   (a) Is it reasonable to use a normal distribution to approximate the number that prefers vinyl? Justify your answer.

   \[
   \mu = 1000(0.25) = 250 \\
   \sigma = \sqrt{1000(0.25)(0.75)} = 13.69 \\
   \text{So yes: normal distribution is ok.}
   \]

   (b) What is the approximate probability that more than 300 people interviewed prefer vinyl?

   \[
   P(X > 300) = P \left( Z > \frac{300 - 250}{13.69} \right) \\
   = P(Z > 3.65)
   \]

   Taking into account that the normal tail only goes up to \( z = 3.5 \), but \( P(Z > 3.5) \approx 0.0003 \), so we can say \( P(X > 300) \approx 0.0003 \), i.e. very small.