Quiz 3 Hints and Comments

Stat 134, Spring 2020, Kolesnik

03/02/2020

These hints and comments are meant to help you fix your work if you missed any questions. They contain common sources of error and hints to push you in the right direction.

Question 1

a. For the mean, you can use the definition of expected value of a random variable. For the variance, just use $E(X^2) - (EX)^2$.
Common errors:

- Some of you tried to use the method of estimators and expressed X as a sum of indicator functions, but X represents just <u>one</u> trial, and therefore one indicator. So, you can use the method of indicators, but as follows

$$E(X) = E(1_A) = P(A) = p$$

where A is the event "success".

- Some of you wrote the expected value of X as an average of n identical probabilities p. This is incorrect. You can think of E(X) as an average, but as an average of all possible values of X (in this case only 0 and 1) weighted by their probabilities.

b. You can exploit the fact that a Binomial random variable X can be seen as a sum of n independent Bernoulli X_i or apply the method of indicators.
Common errors:

- It is not enough to just multiply by n the results from the previous point, you should use linearity of expectation (for the mean) and independence of Bernoulli (for the variance).
- Some of you just stated that $\text{var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \text{var}(X_i)$. But this is true if the X_i's are independent, so you have to point it out.
- For a Binomial, it is not true that $E(X^2) = E(X)$. This statement is correct for a Bernoulli r.v., but not for a sum of Bernoulli r.v., indeed

  $$E(X^2) = E((\sum_{i=1}^n X_i)^2) \neq E(\sum_{i=1}^n X_i) = E(X).$$

  Check it using the method of indicators!

Question 2

a. Sum over the possible values that Y can get.
   Common errors:

   - Summing over Y and then over X, which gives you 1 (why?).
   - Summing over X instead of Y.
   - Providing a different value for the probability of X given different values of Y. That's not the marginal of X!

b. Many of you understood that the idea was showing that we cannot write the joint distribution as product of the marginals.

   To disproof independency, it is enough to check that, for one specific pair of $x, y \in \{1, 2, 3\}$, $P_X(x)P_Y(y) \neq P_{X,Y}(x, y)$. For example, that $P_X(1)P_Y(1) = \frac{1}{4} \neq \frac{1}{16} = P_{X,Y}(1, 1)$.
   Note: proving that, for a specific choice of x and y, $P_X(x)P_Y(y) = P_{X,Y}(x, y)$ is not sufficient to prove independency! You should prove that, <u>for all</u> values $x, y \in \{1, 2, 3\}$, $P_X(x)P_Y(y) = P_{X,Y}(x, y)$.