POL 201.30 – Introduction to Statistical Methods in Political Science
Online Quiz #3– Solution Key

Total Quiz Score:
- **9 points.**

Notes:
- For "**Multiple-Choice**" questions with just one correct answer:
  - The "🟢" symbol next to a bolded option represents the *correct answer*.
- For "**Select All That Apply**" questions:
  - The "✅" symbol next to a bolded option indicates a *correct* selection.
  - The "❎" symbol next to an option indicates an *incorrect* selection.

Question 1 (3 points)

Suppose we conducted a survey of 1,000 individuals. The following information was obtained:

1) 400 of the respondents hold a college degree.
2) 400 respondents voted Republican in the last election and do not hold a college degree.
3) 100 respondents voted Republican in the last election and hold a college degree.
4) 50 respondents did not vote in the last election.
5) 40 respondents both did not vote and did not hold a college degree.
6) All individuals who voted did so for either the Republican or Democratic Party (i.e., no one voted for third parties).

Now, if one individual is selected at random from these 1,000 respondents (that is, each individual is equally likely to be chosen), determine the probability of each of the following events:

*(**Note**: Incorrect or false selections will result in partial credit deductions. Additionally, failure to fully show your work and provide justification for answers will lead to further credit penalties.)*
*(**Tip**: It is strongly recommended that you use a cross-tabulation table in determining your answers and calculations).*

Options:

✅ The probability the respondent did not get a college degree is 60%
✅ The probability the respondent voted Republican is 50%
❎ The probability the respondent voted Democrat is 50%
✅ The probability the respondent voted Republican, given they did not get a college degree is 66.67% app.
❎ The probability the respondent did not get a college degree, given they voted Republican is 2/3.
❎ The probability the respondent voted Democrat, given they do hold a college degree is 29%.

✅ (+0.5 ***if selected***, ***-0.5 if NOT selected***)
❎ (+0.5 ***if NOT selected***, ***-0.5 if selected***)

To solve the question, making a cross-tabulation table using the survey data is crucial.

Essentially to find the missing data, we use the following two properties. Let S be the sample space, let A be an event, A^c its complement, and let B be another event and B^c its complement. Note that S is the universal set, and any event is a subset of S. Then,

1) **Rule #1: The sample space will always equal the union of an event and its complement, i.e., $S = A \cup A^c$.**
2) **Rule #2: Any event can be broken down as the union of i) the event itself joined with another event, and ii) the event itself joined with the complement of the other event. That is, $A = (A \cap B) \cup (A \cap B^c)$.**

These two rules allow us to infer the missing components of the table.

The first step in creating the table is looking at which variables interact, such as the 400 who voted Republican and do not hold a college degree. Because of this statement, one side of the table will list who the individual voted for (*columns*) and the other whether they attended college (*rows*).

- Next, since we know the total number of respondents is 1000, the total number of those who have *attended college* **or** *didn't attend* sum up **1000**.
- We also know that no third party received a vote. Thus, the count for respondents who voted *Republican*, *Democrat*, **or** *Did Not Vote* sums to **1000**.

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)				
No College (NC)				
Column Total				**1000**

- Next, the first two bullets' state, "***400*** *of the respondents hold a college degree*" and "***400*** *respondents voted Republican in the last election and do not hold a college degree*"
  - Now we can fill in these two boxes:

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)				**400**
No College (NC)	**400**			
Column Total				1000

- Then from this we can fill in the *No College* total of 600 because 1000 = 400 + #(*No College*); thus #(*No College*) = 1000 – 400 = **600**.
  - Here we are applying the following property of the **sample space** set: "the sample space is always equal to the union of an event A and its complement A^c," or $S = A \cup A^c$ (in this context, S is also the universal set).
- Moreover, the third bullet "***100*** *respondents voted Republican in the last election and hold a college degree*" allows us to directly input this number in the table:

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)	***100***			400
No College (NC)	400			**600**
Column Total				1000

- The total number of Republicans can be calculated by adding the **100** respondents with *College* and 400 with *No College*, which totals 500.
    - This is supported by the following rule. Say you have an event R, representing "*A respondent voted Republican*." Then, you can always break it down further by taking its intersection with another event E ("*Attended College*") and its complement E^c ("*Didn't Attend College*") and then taking their union. That is,

      $$\boldsymbol{R = (R \cap E) \cup (R \cap E^c)}$$

      Given that $(R \cap E)$ and $(R \cap E^c)$ are *mutually exclusive events*, we can conclude:

      $$Total\ Elements\ in\ R = \#R = \#(R \cap E) + \#(R \cap E^c)$$

      Note that this is the basis for the law of total probability: $P(R) = P(R \cap E) + P(R \cap E^c)$. If you need to further convince yourself this is correct, do a Venn Diagram.

- The final two bullets are "**<u>50</u>** respondents did not vote in the last election" and "**<u>40</u>** respondents both did not vote and did not hold a college degree".

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)	**100**			400
No College (NC)	400		**<u>40</u>**	600
Column Total	**500**		**<u>50</u>**	1000

- From here we can fill in the remaining boxes. In every case, we are implicitly using a Venn Diagram and the two mentioned **properties of events and the sample space** to find the missing pieces.
    - First the total number of respondents who voted Democrats can be found as follows.
        - Since there are 1000 total respondents and the respondents that voted *Republican* is 500 and those who *Did Not Vote* is 50, **the number of Democrats equals the difference (Rule #1).**
            - This equation is 1000 - 500 - 50 = **450**

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)	**100**			400
No College (NC)	400		**<u>40</u>**	600
Column Total	**500**	**450**	**<u>50</u>**	1000

-
    - Next, the number of respondents who voted Democrat and have no college can be found as follows.
        - Since the total of No College respondents is 600, and we have the number of Republicans and Did Not Voter respondents who have no college, we can find the number of Democrat voters (**Rule #2):**
            - This equation is 600 - 400 - 40 = **160**

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)	**100**			400
No College (NC)	400	**160**	**<u>40</u>**	600
Column Total	**500**	**450**	**<u>50</u>**	1000

-
    - Then, the total number of respondents who voted Democrat can be found by subtracting the number of no college respondents of 160 from the total number of 450 (**Rule #2**):
        - This equation is 450 - 160 = **290**
    - And finally, we can do the simply subtract the number of no college and democrat voting respondents, which is 40, from the of the total number Did Not Vote, which is 50, to find the respondents who went to college and did not vote (**Rule #2**):
        - 50 - 40 = **10**

- Here is the completed table:

	Republican (R)	*Democrat (D)*	*Did Not Vote (NV)*	Row Total
College (C)	100	**290**	**10**	400
No College (NC)	400	**160**	40	600
Column Total	500	**450**	50	1000

Options:

- ✅ **1) The probability the respondent did not get a college degree is 60%**
- ✅ **2) The probability the respondent voted Republican is 50%**
- ❌ 3) The probability the respondent voted Democrat is 50%
- ✅ **4) The probability the respondent voted Republican, given they did not get a college degree is 66.67% app.**
- ❌ 5) The probability the respondent did not get a college degree, given they voted Republican is 2/3.
- ❌ 6) The probability the respondent voted Democrat, given they do hold a college degree is 29%.

Justification of Correct Selections:

- **1st option**: **Correct.** The probability that the selected respondent did not get a college degree is 60%. If we look at the table we created, 600 of the 1000 total respondents did not get a college degree. 600 of 1000 is 60%.
- **2nd option**: **Correct.** The probability that the selected respondent voted Republican is 50%. Again, looking at the table, 500 of the total 1000 respondents voted Republican. 500 of 1000 is 50%.
- **3rd option**: **Incorrect.** The probability that the selected respondent voted Democrat is not 50%. The number of respondents that voted Democrat was 450. Out of the total 1000 respondents, Democrats accounted for 45% of the total respondents because 450/1000 is 0.45 or 45%, not 50%
- **4th option**: **Correct.** The probability that the selected respondent voted Republican, given they did not get a college degree is 66.67%. The number of respondents that voted Republican and have no college degree is 400. The total number of no college respondents was 600. To find the probability of selecting a respondent that voted Republican and do not have a college degree, *you must find the percentage of the total no college respondents who voted Republican.* This would be 400/600 which equals 0.6667 or 66.67%.
- **5th option**: **Incorrect.** The probability that the selected respondent did not get a college degree, given they voted Republican is not 2/3. The number of respondents that voted Republican was 500 and the number of these 500 that do not have a college degree is 400. The probability of selecting a respondent that has no college degree from the Republican voters is 400/500 or 4/5, not 2/3.
- **6th option**: **Incorrect.** The probability that the selected respondent voted Democrat, given they do hold a college degree is not 29%. The total number respondents that hold a college degree is 400 and the number of these 400 that voted Democrat is 290. The probability would be the number of Democrat-voting, college- educated of 290 divided by the total college degree respondents of 400. This would be 290/400 or 0.725 or 72.5%, not 29%.

Question 2 (2 points)

Two friends are discussing the probability of winning a game. One says, "*I calculate there is a 90% probability that I will win on my next turn. Therefore, what happens next is not really random.*"

Select the correct statement:

Options:

- [x] The friend is wrong because, according to the definition of a random process, an event can still be random even if one outcome is much more likely than others.
- [ ] The friend is right because a random process requires that all possible outcomes occur with equal probability.
- [ ] The friend is wrong because probability only applies to events that have already happened, not future outcomes.
- [ ] The friend is right because an event with a 90% probability is effectively predetermined.

- [x] *(+2 points)*
- [ ] *(0 points)*

Justification:

- The student is incorrect because the probability that that they will win on the next term is random. The definition of a random process is one where the outcome is uncertain before it happens. Since the next turn has not taken place yet, it is uncertain what will happen, all outcomes are still possible. Even though the student predicts they have a 90% that they will win on the next term, it is not certain. What happens on the next turn is random and not known. This describes why the first answer choice is correct, because it is a random process. This eliminates both answer choices where it notes the student is correct. The third answer choice is also incorrect because random process applies to future outcomes, not past outcomes.

Question 3 (1 points)

Imagine you are flipping a fair coin (that is, the probability of landing heads is 0.50). You have flipped the coin 15 times, and 14 of those flips have landed on tails.

Select the correct statement:

Options:

- [x] The probability that the 16^{th} flip lands on tails is 0.5.
- [ ] The probability that the 16th flip lands on tails is higher than the probability of landing on heads.
- [ ] The probabilities of heads and tails are the same, but it is impossible to determine their exact value without more information.
- [ ] The probability that the 16^{th} flip lands on heads is higher than the probability of landing on tails.

- [x] *(+1 point)*
- [ ] *(0 points)*

Justification:

- The correct answer choice is "The probability that the 16^{th} flip lands on tails is 0.5" because each time a coin is flipped the probability of landing tails is 0.50. *Each coin flip is its own independent event* that is not influenced by past flips. Since there is only two possible outcomes with equal probability, heads or tails, each side has a 50% of landing. We know the exact probability of the coin landing on tails on the 16^{th} flip and it is a 0.5.

Question 4 (2 points)

Which of the following statements correctly describes disjoint (mutually exclusive) events?

<u>Options</u>:

- (●) Two events that cannot happen at the same time.
- ( ) Two events that always have the same probability.
- ( ) Two events that can occur together, but with a reduced probability.
- ( ) Two events that are always independent.

(●) (*+2 point*)
( ) (*0 points*)

Justification:

- The correct choice is "*Two events that cannot happen at the same time*" because, by definition, disjoint mutually exclusive events cannot happen at the same time. The definition of "mutually exclusive" from the Cambridge Dictionary is that "if two things are mutually exclusive, they cannot exist or happen together at the same time". This also can be determined from our lecture on probability, slide 12, where is describes *Disjoint (Mutually Exclusive) Events as two events are disjoint if they do not share any outcomes*. For example, if you have a die, it is not possible to roll a 3 and roll a 5 at the same time. They are mutually exclusive.

Question 5 (1 points)

Which of the following best defines a random process?

<u>Options</u>:

- (●) A process where the outcome is uncertain before it happens.
- ( ) A process where all outcomes occur with equal probability.
- ( ) A process where the same initial conditions always lead to the same outcome.
- ( ) A process where the probability of each outcome is not known.

(●) (*+1 point*)
( ) (*0 points*)

Justification:

- The correct choice is "*A process where the outcome is uncertain before it happens*" because it is most similar to the definition we have learned in class. The definition for random process is given in our lecture on probability on slide 6. The definition reads "*a random process involves initial conditions leading to an uncertain outcome*". The final outcome is unknown prior to the process happening. An example of a random process is an election. The initial conditions are the current political climate, economic conditions, and media coverage, the outcome is the election result or who won. There is uncertainty of who will win the election prior to election day, hence making it a random process.