

INSTRUCTOR- PROFESSOR BRAD LEVECK

GRADUATE STUDENT – RUTH CHIKA UZOKA

COURSE CODE AND TITLE- POLI 210(QUANTITATIVE ANALYSIS OF POLITICAL DATA 1)

HOMEWORK – A5

ANSWERS-

- 1- Research Question: What is the extent of the influence of social media usage on the outcomes of political campaigns?

Assumptions:

- 1 -Voter's views are significantly shaped by exposure to political content on social media.
- 2 -Social media is a prominently utilized platform for political campaigns, facilitating social communication and public engagement.

Hypotheses:

- 1 -Effectively utilizing social media for political campaigns will result in increased awareness and higher levels of political engagement.
- 2 -Political campaigns will have a more pronounced impact on the opinions of younger voters who are actively engaged on social media platforms.

2- The relative frequency distribution distinguishes itself from a frequency distribution by presenting the proportion or percentage of observations within each category, rather than merely providing the raw count of observations in each category. In essence, it focuses on the relative proportions or percentages to convey the distribution's characteristics.

3- It is impossible to create a Probability density function (PDF) by plotting the graph of the relative frequency distribution of each value in the sample because while PDF reflects the series of the probability distribution of a random variable, the relative frequency distribution is based on discrete observations.

4- The distinction between a Probability Mass Function (PMF) and a Cumulative Distribution Function (CDF) lies in their purposes: a PMF is employed to depict the probability distribution of a discrete random variable, while a CDF provides the probability that a random variable assumes a value less than or equal to a specified value. In essence, the PMF characterizes the likelihood of each discrete outcome, while the CDF accumulates these probabilities up to a given threshold value.

5- A contingency table is a valuable tool when investigating the interconnection of two variables, especially when examining the relationship between categorical variables. For example, it can be effectively utilized to scrutinize how gender and voting preferences are related and influence each other.

6- Two political processes that could be represented by discrete distributions include:

Bernoulli or Binomial Distribution: This type of distribution is suitable for modeling the results of single elections where there are only two possible outcomes, such as whether a Democrat or Republican would win or lose.

Poisson or Negative Binomial Distribution: These distributions are often employed to model the occurrence of political protests or demonstrations within a specific time frame, providing insights into their frequency and distribution.

7-

8-

9- With a mean of 3 events occurring in a year, we are interested in calculating the probability of having 4 events in a year.

Given that the natural logarithm of base e = \ln base e which is = 2.71828.

Therefore, $P(X = 4) = \frac{(-3)^4}{4!}$

$= \frac{2.71828^4 \cdot (-3)^4}{4!}$

$= \frac{4.028}{24}$

$= 0.168$

Hence, the probability that there would be 4 wars in any given year is 0.168

10 - Given that;

(Expected Utility) = (Probability of Event 1 * Utility of Event 1) + (Probability of Event 2 * Utility of Event 2)

The probability of being at Bach's concert is 0.5

Bach = $18 + 8 = 26$

The probability of being at the Stravinsky concert is 0.5

$= 3 + 8 = 11$

Expected Utility = $(0.5 * 26) + (0.5 * 11) = 13 + 5.5 = 18.5$

Therefore,

Bach concert when probability is = 0.3

$18 + 8 = 26$

at the Stravinsky concert when probability = 0.7

$= 3 + 8 = 11$

Expected Utility = $(0.3 * 26) + (0.7 * 11) = 7.8 + 7.7 = 15.5$

Since her expected utility is higher when she goes for BACH, (18.5) She should go to the Bach concert.

11- Utility (Start a War) = (Probability of Winning * Utility of Winning) + (Probability of Losing * Utility of Losing)

For Nation B, probability given when weak and strong;

probability given strong = $(0.60 * 1) + (0.40 * (-1)) = 0.60 - 0.40 = 0.20$

Probability given weak = $(0.30 * 1) + (0.70 * (-1)) = 0.30 - 0.70 = -0.40$

Overall outcome =

Utility (Start a War) = (P of B being Strong * Expected Utility | Strong B) + (P of B being Weak * Expected Utility | Weak B)

Therefore,

Utility = (Start a War) = $(0.40 * 0.20) + (0.60 * (-0.40)) = 0.08 - 0.24 = -0.16$

Given the calculated expected utilities, it is in Nation A's best interest not to initiate a war after observing the maneuvers. The expected utility of refraining from war (0) outweighs the expected utility of commencing a war (-0.16).

EXERCISES 11.4-

1- The reason why creating a PDF by plotting the graph of the relative frequency distribution is not feasible lies in the nature of the relative frequency distribution. It offers information about the frequencies of distinct values within the sample, but it lacks the continuous probability density function necessary for generating a PDF.

2- Probability mass functions (PMFs) are employed to characterize discrete probability distributions, while probability density functions (PDFs) are utilized for representing continuous probability distributions.

3- A PDF provides the probability of a random variable assuming a precise value identical to that variable, whereas a CDF offers the probability of a random variable assuming a value less than or equal to the variable.

4- A scatter plot proves invaluable for investigating the joint distribution relationship between two variables when our objective is to comprehend their connection. In political science, for instance, we might employ a scatter plot to analyze the correlation between campaign spending and electoral outcomes in a particular election.

5- Assessing whether a probability distribution is normal by visual inspection alone is not reliable, as the distribution's shape can be deceptive. It might appear normal but actually follows a different distribution with a similar outward appearance. To establish the true nature of the distribution, we must rely on statistical tests or employ specialized data analysis techniques.

6- The relative frequency distribution graph visually represents the proportion or percentage of observations within each interval or category of a variable. On the other hand, a probability density

function (PDF) is a mathematical function that quantifies the likelihood of a random variable assuming values within a particular range.

7- To compute probabilities using a probability density function (PDF), an integral calculation is necessary since it corresponds to the area beneath the PDF curve. Integrating the PDF across a specified range enables us to determine the probability of the random variable falling within that range. This integral process is crucial for calculating probabilities and conducting various statistical analyses, leveraging the properties inherent to the PDF.

$$\begin{aligned} 8- \text{show that } \text{VAR}(X) &= E[(X-u)^2] = E[X^2] - u^2 \\ &= (X^2 - 2Xu + u^2) \\ &= X^2 - 2EXu + u^2 \\ &= X^2 - 2u^2 + u^2 \\ &= E(X^2) - u^2 \end{aligned}$$

$$9 - X > 10 = [25 - 10] / [25 - 0] = 15/25$$

$$= 0.6 \text{ or } 60\%$$

10 – The mean budget deficit of \$100 million, along with a standard deviation of \$300 million, provides insights into the country's financial situation. The standard deviation indicates that the budget deficits vary around the mean by an average of \$300 million. This reveals a significant range in budget deficits, with some exceeding the mean while others fall below it.

11- **Logistic Distribution** - The distribution of voter turnout in a particular election. This follows a logistic distribution since the voter turnout may vary, but it is likely to be clustered around a certain average value, with a clear upper and lower limit.

Exponential Distribution - The distribution of time between political scandals. This distribution may have a long tail, indicating that most scandals occur relatively close together in time, but there are occasional long periods without scandals.

Normal or Log-Normal Distribution - The distribution of campaign contributions from individuals to political candidates. This follows a normal or log-normal distribution because the amount contributed may vary, but it is likely to be clustered around a certain average value, with most contributions falling close to this average.