Vanderbilt University
Political Science Department
Fall 2023

Stats 1
(PSCI 8356)
Professor Jim Bisbee

PROBLEM SET 2: Due Tuesday, September 26th at beginning of class.

A reminder: you may work with others in the class on this problem set, and you are in fact encouraged to do so. However, the work you hand in must be your own. Handwritten work is acceptable, but word-processed work (e.g., using LaTeX or `RMarkdown`*) is preferred.*

1. A single unfair die is tossed once. Let Y be the number facing up. Find the expected value and variance of Y given the following probabilities for each face: $P(s_1) = 0.2$, $P(s_2) = 0.1$, $P(s_3) = 0.15$, $P(s_4) = 0.15$, $P(s_5) = 0.2$, $P(s_6) = 0.2$.

2. Consider a discrete random variable X with a mean μ and variance σ^2. Now define a new random variable $Y = X + 11$.

   (a) Is the mean of Y greater than, less than, or equal to the mean of X? Show your work using Theorems 3.3 and 3.5 from WMS.

   (b) Building on the answer to the preceding question, is the variance of Y greater than, less than, or equal to the variance of X?

   (c) Now consider a new random variable $Y = 2X$. Redo the analyses in 3.a and 3.b to determine if the mean and variable of Y is different from X.

   (d) Again, re-run the same analyses for a new random variable $Y = X/10$.

   (e) Finally, calculate the mean and variance of a new random variable $Y = aX + b$, where a and b are *constants*.

3. Djokovic and Medvedev play a series of games until one player wins three games. We assume that the games are played independently and that the probability that Djokovic wins any game is p. Compute the probability that the series lasts exactly five games. [Hint: Use what you know about the random variable, Y , the number of games that Djokovic wins among the first four games.]

4. A CCTV camera is placed above an un-signed (i.e., no stop signs or stop lights) intersection on a quiet dirt road in Vermont and records for one month. In that time, there are 100 instances of two cars arriving at the intersection at the same time, and three accidents.

   (a) What is the probability of observing at least 5 accidents in the next month, assuming that the total number of instances with two cars remains at 100?

   (b) If you observe more than 5 accidents in the next month, would you conclude that the rate of 3 accidents per 100 instances has changed? Explain.

5. Stats 1 students are given one hour to complete their midterm exam. The amount of time that they actually take is a continuous random variable described with a probability density function:

$$f(y) = \begin{cases} cy^2 + y & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Solve for c. (HINT: remember axiom 2!)
   (b) Find $F(y)$ (i.e., the CDF)
   (c) Using the CDF function, calculate the probability that a randomly selected student finishes immediately or "faster" (i.e., $P(Y \le 0)$).
   (d) Using the CDF function, calculate the probability that a randomly selected student finishes within an hour (i.e., $(P(Y \le 1))$.
   (e) Calculate the probability that a randomly selected student finishes within 30 minutes.
   (f) Calculate the probability that a randomly selected finishes between 15 minutes and 45 minutes.
   (g) Given that a particular student needs at least fifteen minutes to finish the exam, calculate the probability that she will need *at least* 30 minutes to finish. (NB: remember the definition of conditional probability, and pay attention to the words "at least"!)

6. Upon studying low bids for shipping contracts, a microcomputer manufacturing company finds that intrastate contracts have low bids that are uniformly distributed between 20 and 25, in units of thousands of dollars. Find the probability that the low bid on the next intrastate shipping contract:

   (a) is below $22,000
   (b) is above $24,000

7. Using `R`'s 'pnorm()' function, calculate the following for a standard normal random variable Z:

   (a) $P(0 \le Z \le 1.2)$
   (b) $P(-.2 \le Z \le .2)$
   (c) $P(-1.56 \le Z \le -.2)$

8. Using `R`'s `qnorm()` function, find the value z_0 such that:

   (a) $P(Z \ge z_0) = 0.5$
   (b) $P(Z \le z_0) = 0.8643$
   (c) $P(-z_0 \le Z \le z_0) = 0.90$
   (d) $P(-z_0 \le Z \le z_0) = 0.95$

[THERE IS ONE MORE PROBLEM ON THE NEXT PAGE.]

9. **EXTRA CREDIT:** Consider the following standard setup in formal models of electoral competition, in which the utility voter v derives from electing candidate a is written

$$\begin{aligned} U_v(x_a) &= -(x_v - x_a)^2 \text{, where} \\ & \quad x_v \text{ is the voter's ideal policy on the real line, and} \\ & \quad x_a \text{ is the policy (also on the real line) candidate } a \text{ will enact if elected.} \end{aligned}$$

(A concrete way to think about this, for example, is to consider x_a and x_v to be two different tax rates.)

(a) [Easy; not a trick question.] Say that a makes a binding proposal during an election campaign to enact x_a if elected. What proposal (and therefore what policy) maximizes v's utility?

Now consider the case where the voter is unsure about what a will do if elected. A reasonable way to model this scenario would be to consider x_a a random variable with mean μ_a and variance σ_a^2. In this case, rather than evaluating $U_v(x_a)$, the voter evaluates her expected utility, or $E[U_v(x_a)]$.

(b) Supply an expression for $E[U_v(x_a)]$ written only in terms of x_v, μ_a, and σ_a^2.

(c) What is $\frac{\partial E[U_v(x_a)]}{\partial \sigma_a^2}$?

(d) Have we made any assumptions about the distribution of x_a?

(e) As specified above, $U_v(x_a)$ is an example of what is known as a "concave utility function." Based on your analysis here, why is it appropriate that agents with concave utility functions are said to be "risk averse?"

(f) If we assume that voters are risk averse, do candidates have an incentive to be vague in a campaign about the policies they'll enact if elected? (HINT: Your response should explicitly refer to how σ_a^2 affects $E[U_v(x_a)]$.)