Economics 30330: Statistics for Economics
Problem Set 3 - Suggested Answers
University of Notre Dame
Instructor: Julio Garín
Spring 2012

Basics of Probability (100 points)

1. Bruno's offers pizzas with two different sizes, three different types of crusts, and up to 10 toppings. How many different types of pizzas can be made?

   When you order a pizza, you must choose a size, a crust, and whether or not you have each of the toppings. The first two parts are easy: you can pick one of two sizes, C_1^2*, and one of three crusts,* C_1^3*.*

   The problem is knowing how to model the number of topping combinations. For each topping, you also have two decisions: whether to order that particular topping, or not. Every pizza will come with or without pepperoni, with or without onions, with or without sausage, etc. As a result, for each topping, there are C_1^2 *choices, or* 2^{10} *choices in all.*

   In the end, there are a total of $C_1^2 C_1^3 (C_1^2)^{10} = (2)(3)(2^{10}) = 6144$ *possible types of pizzas.*

2. Suppose that two runners from team A and three runners from team B participate in a race. If all five runners have equal ability and there are no ties, what is the probability that two runners from team A will finish first and second, and that the three runners from team B will finish third, fourth, and fifth?

   The possible arrangements in which the 5 runners can finish the race with no ties is $P_5^5 = 5!$*. If the two runners from team A finish in the first two positions, there are* $P_2^2 = 2!$ *arrangements of these two runners among these two positions, and* $P_3^3 = 3!$ *arrangements of the three runners from team B among the last three positions. Therefore, there are 2!3! arrangements in which the runners from team A finish in the first two positions and the runners from team B finish in the last three positions. Let E be the event that the runners from team A finish in the first two positions and the runners from team B finish in the last three positions.*

   P(E)=(# ways the runners from team A finish in the first two positions and the runners from team B finish in the last three positions)/(Total # of possible arrangements). Therefore,

   $$P(E) = \frac{2!3!}{5!} = 0.1$$

3. A high school senior applies for admission to the University of Northern Detroit and University of Southern Colorado. She estimates the probability of acceptance at ND at 0.8, the probability of acceptance at USC at 0.3, and the probability of being be admitted to both at 0.2. What is the chance she will not be accepted at either school?

   Note that P(Admitted to neither school) and P(Admitted to at least one school) are complements. Therefore, P(Admitted to neither school)=1-P(Admitted to at least one school).

   $P(ND \cup USC) = P(ND) + P(USC) - P(ND \cap USC) = 0.8 + 0.3 - 0.2 = 0.9$. *Therefore,*

   $$\begin{aligned} P(ND^C \cap USC^C) &= 1 - P(ND \cup USC) \\ &= 1 - 0.9 \\ &= 0.1 \end{aligned}$$

4. Consider the experiment of rolling a pair of dice.

   (a) How many sample points are possible?
   $(6)(6) = 36$

   (b) List the sample points.

$$\begin{array}{ccccc} 1,1 & 2,1 & 3,1 & \cdots & 6,1 \\ 1,2 & 2,2 & 3,2 & \cdots & 6,2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1,6 & 2,6 & 3,6 & \cdots & 6,6 \end{array}$$

   (c) What is the probability of rolling a 7?

$$P(S = 7) = \frac{1}{6}$$

   (d) What is the probability of rolling a 9 or greater?

$$P(S \geq 9) = \frac{P(S = 9) + P(S = 10) + P(S = 11) + P(S = 12)}{36} = \frac{4 + 3 + 2 + 1}{36} = \frac{5}{18}$$

   (e) What is the probability of rolling an odd number?

$$P(odd) = \frac{18}{36} = \frac{1}{2}$$

   (f) What is the probability of rolling an even number?

$$P(even) = 1 - P(odd) = \frac{1}{2}$$

5. Listed below is a two-by-two table. In period 1, events A or B can happen. In period 2, outcome C or D will result. Assume $\Pr(C|B) = 0.15$ and $\Pr(D|A) = 0.7$. Please fill in the missing boxes below.

	Outcome A	**Outcome B**	
Outcome C	0.24	0.03	0.27
Outcome D	0.56	0.17	0.73
	0.80	0.20	1.00

Use the definitions of conditional probabilities:

$P(C|B) = \frac{P(C \cap B)}{P(B)}$ *so* $P(B) = \frac{P(C \cap B)}{P(C|B)} = \frac{0.03}{0.15} = 0.2$

Since $P(A) + P(B) = 1$, *then* $P(A) = 1 - P(B) = 1 - 0.2 = 0.8$.

By definition, $P(B) = P(C \cap B) + P(D \cap B)$.
Therefore, $P(D \cap B) = P(B) - P(C \cap B) = 0.20 - 0.03 = 0.17$.

You are given that $P(D|A) = 0.7$.
Therefore, $P(D|A) = P(D \cap A)/P(A)$ *so* $P(D \cap A) = P(D|A)P(A) = (0.7)(0.8) = 0.56$.

$P(A) = P(C \cap A) + P(D \cap A)$. *So* $P(C \cap A) = P(A) - P(D \cap A) = 0.8 - 0.56 = 0.24$.

6. Small cars may not be as safe as large cars. Small cars accounted for 18% of all vehicles on the road. Accidents involving small cars led to 11,898 fatalities during a recent year. Assume the probability that a small car is involved in an accident is 0.18, the probability that an accident involving a small car leads to a fatality is 0.128. Similarly, the probability that the accident leads to a fatality driving a car that is not small is 5%. You hear of an accident involving a fatality. Assume that the probability of getting into an accident is independent of car size. What is the probability that a small car was involved?

   $S = \textit{Small Car}$
   $F = \textit{Fatality}$

   $P(S) = 0.18$
   $P(S^c) = 0.82$
   $P(F|S) = 0.128$
   $P(F|S^c) = 0.05$

   $$\textit{We need to find } P(S|F) = \frac{P(F|S)P(S)}{P(F)}$$

   $$P(F) = P(F|S)P(S) + P(F|S^c)P(S^c)$$

   $$P(S|F) = \frac{(0.18)(0.128)}{(0.128)(0.18) + (0.05)(0.82)} = \frac{0.023}{0.064} = 0.36$$

7. The prior probabilities for events A_1, A_2 and A_3 are $\Pr(A_1) = 0.2$, $\Pr(A_2) = 0.5$, and $\Pr(A_3) = 0.3$. The conditional probabilities of B given A_1, A_2 and A_3 are $\Pr(B|A_1) = 0.5$, $\Pr(B|A_2) = 0.4$, and $\Pr(B|A_3) = 0.3$.

   (a) Find $\Pr(B \cap A_1), \Pr(B \cap A_2)$ and $\Pr(B \cap A_3)$.

   $$Pr(B \cap A_1) = Pr(A_1)Pr(B|A_1) = (0.2)(0.5) = 0.1$$

   $$Pr(B \cap A_2) = Pr(A_2)Pr(B|A_2) = (0.5)(0.4) = 0.2$$

   $$Pr(B \cap A_3) = Pr(A_3)Pr(B|A_3) = (0.3)(0.3) = 0.09$$

   (b) Apply Bayes' theorem to compute the posterior probability $\Pr(A_2|B)$.

   $$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}$$

   $$= \frac{P(B \cap A_2)}{P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3)} = \frac{0.2}{0.1 + 0.2 + 0.09} = 0.51$$

   (c) Use the tabular approach to applying Bayes' theorem to compute $\Pr(A_1|B), \Pr(A_2|B)$, and $\Pr(A_3|B)$.

Event	$P(A_i)$	$P(B\|A_i)$	$P(A_i)P(B\|A_i)$	$P(A_i\|B_i)$
A_1	0.2	0.5	0.1	0.1/0.39=0.26
A_2	0.5	0.4	0.2	0.2/0.39=0.51
A_3	0.3	0.3	0.09	0.09/0.39=0.23
			0.39	

8. According to the Arizona Chapter of the American Lung Association, 7% of the population has lung disease. Of those having lung disease, 90% are smokers; of those not having lung disease, 25.3% are smokers. What is the probability that a randomly selected smoker has lung disease?

   Let S = event that person selected is a smoker

   Let L = event the person selected has lung disease.

   Let L^C = event the person selected does not have lung disease.

   L and L^C are complementary $\Longrightarrow$ mutually exclusive and exhaustive.

   $\Pr(L) = 0.07$
   $\Pr(S|L) = 0.90$
   $\Pr(S|L^C) = 0.253$

   Since L^C= not L, then $\Pr(L)$=*1*-$\Pr(L^C) = 1 - 0.07 = 0.93$.

   Apply Bayes' theorem:

$$\Pr(L|S) = \frac{\Pr(L)Pr(S|L)}{Pr(L)Pr(S|L) + Pr(L^C)Pr(S|L^C)} = \frac{(0.07)(0.90)}{(0.07)(0.90) + (0.93)(0.253)} = 0.2112$$

9. Prove the following:

   (a) The mean and variance of the z-score are 0 and 1, respectively.

   First the mean:

$$\begin{aligned}\bar{z} = \frac{\sum_i z_i}{n} &= \frac{\sum_i \left(\frac{x_i - \bar{x}}{s}\right)}{n} \\ &= \frac{\sum_i (x_i - \bar{x})}{sn} = \frac{0}{sn} \\ &= 0\end{aligned}$$

   Now the variance:

$$\begin{aligned}s_z^2 &= \frac{\sum_i (z_i - \bar{z})^2}{n-1} = \frac{\sum_i z_i^2}{n-1} \\ &= \frac{\sum_i \left(\frac{x_i - \bar{x}}{s}\right)^2}{n-1} \\ &= \frac{\sum_i (x_i - \bar{x})^2}{s^2(n-1)} \\ &= \frac{1}{s^2}\frac{\sum_i (x_i - \bar{x})^2}{n-1} = \frac{1}{s^2}s^2 \\ &= 1\end{aligned}$$

   Note that you may also show this using the population mean and variance.

   (b) $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = \sum_{i=1}^{n} x_i - n\bar{x} = \sum_{i=1}^{n} x_i - n\frac{\sum_{i=1}^{n} x_i}{n} = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i = 0$$

10. Recall the formula for population variance when completing the following exercises.

(a) Show that the sum of squared deviations can be written as $\sum_{i=1}^{N} x_i^2 - N\mu^2$.

$$\begin{aligned}\sum_{i=1}^{N}(x_i-\mu)^2 &= \sum_{i=1}^{N}(x_i^2 - 2x_i\mu + \mu^2)\\ &= \sum_{i=1}^{N} x_i^2 - 2\mu\sum_{i=1}^{N} x_i + \sum_{i=1}^{N}\mu^2\end{aligned}$$

Multiply the second term by $\frac{N}{N}$*:*

$$\begin{aligned}&= \sum_{i=1}^{N} x_i^2 - 2\mu\frac{N}{N}\sum_{i=1}^{N} x_i + N\mu^2\\ &= \sum_{i=1}^{N} x_i^2 - 2N\mu^2 + N\mu^2\\ &= \sum_{i=1}^{N} x_i^2 - N\mu^2\end{aligned}$$

(b) Show that the variance can be written as $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \mu^2$.

$$\begin{aligned}\sigma^2 &= \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2\\ &= \frac{1}{N}\left(\sum_{i=1}^{N}(x_i^2 - 2x_i\mu + \mu^2)\right)\\ &= \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2 - 2\mu\sum_{i=1}^{N} x_i + N\mu^2\right)\\ &= \frac{1}{N}\sum_{i=1}^{N} x_i^2 - 2\mu^2 + \mu^2\\ &= \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \mu^2\end{aligned}$$

(c) Show that the variance can also be written as $\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} x_i (x_i - \mu)$.

$$
\begin{aligned}
\sigma^2 &= \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 \\
&= \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu) \\
&= \frac{1}{N} \sum_{i=1}^{N} [x_i(x_i - \mu) - \mu(x_i - \mu)] \\
&= \frac{1}{N} \sum_{i=1}^{N} x_i(x_i - \mu) - \frac{\mu}{N} \sum_{i=1}^{N} (x_i - \mu)
\end{aligned}
$$

Note that we have already proved that $\sum_{i=1}^{N} (x_i - \mu) = 0$, *therefore,*

$$
\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} x_i(x_i - \mu)
$$