Math 102 / Core 143, Sections AX and BX — Exam II

Show all work clearly for partial credit; in particular, write down any difficult computations before you perform them. An unevaluated expression is worth more than the numerical answer without explanation.

1. *(20 points)* A deck of cards for the game of pinochle has only 6 ranks, 9 through ace, but two cards of each rank and suit. So, for example, there are two aces of spades, and there are a total of only 48 cards in the deck. Suppose two cards from a pinochle deck are dealt face-down on the table.

   (a) What is the probability that the first card is an ace?

   (b) Given that the first card is a king, what is the probability that the second is also a king?

   (c) What is the probability that both cards are kings?

   (d) What is the probability that the first card is either a queen or a heart?

   (e) Suppose that one card is dealt face up, noted and returned to the deck, and the deck is reshuffled, all a total of 5 times. What is the probability that at least one of the 5 cards noted is a spade?

2. *(8 points)* Suppose 20 children are paired off at random into 10 pairs; one of each pair is taught to read using the phonics method, the other by the whole-word method. After 10 weeks of instruction, all the children take a standardized reading test. Assume the two teaching methods are equally effective. What is the probability that in at least 8 out of the 10 pairs, the child who was taught by the whole-word method gets a higher score on the test, just by chance?

3. *(15 points)* It is possible to make a solid with eight sides, each an equilateral triangle. Suppose that the sides of such a solid are numbered 1 through 8, to form an 8-sided die, and that this die is rolled a large number of times. Which of the following statements are likely to be true as the number of rolls increases?

   (a) The (absolute value of the) difference between the number of 4's rolled and one-eighth of the total number of rolls decreases.

   (b) The (absolute value of the) difference between the fraction of rolls that are 4's and the number one-eighth decreases.

   (c) The probability histogram for the sum of the rolls more closely approximates the normal curve (when converted to standard units).

   (d) The probability histogram for the average of the rolls more closely approximates the normal curve (when converted to standard units).

   (e) The histogram for the numbers rolled more closely approximates the normal curve (when converted to standard units).

4. *(15 points)* The game of Essel has two variations. Playing the variation Ess, you have 1 out of 5 chances to win and the game pays 3 to 1. Playing the variation El, you have 1 out of 10 chances to win and the game pays 7 to 1. Suppose you play the game 100 times, betting $1 each time.

   (a) If you always play variation Ess, you should expect to win a total of $_____$, give or take $_____$.

(b) If you always play variation El, you should expect to win a total of $_____$, give or take $_____$.

(c) You are more likely to break even if you always play variation _____ (Ess, El or neither).

5. *(18 points)* A marketing survey collects responses by 400 households chosen at random from a city with 100,000 households to learn average household income. The average household income in the survey is $24,000 with a standard deviation of $10,000.

(a) This survey process can be modelled as drawing tickets from a box. How many draws?

(b) When estimating the average income, what values should the tickets in the box be? Choose one answer:

income levels   ones and zeros

(c) True or False: We don't know the standard deviation of household income for the city, but we can approximate it with the standard deviation of the incomes in the sample.

6. *(16 points)* In the survey described in question 5 above (again, with a sample of 400 households), additional data is collected:

• The number of households with at least one pet is 180.

• The number of people per household averages 2.4 with a standard deviation of 1.4.

• The average for the 400 households in the survey spent on pet food each month is $54 with a standard deviation of $20.

Answer the following questions, if possible, using the above information. If it is not possible given the data above, write the formula you would use and identify what additional information is needed.

(a) Estimate (with error estimate) the percentage of households in the city which have pets.

(b) Estimate (with error estimate) the number of households in the city with more than 2 people.

(c) Give a 95% confidence interval for the average amount each household in the city spends on pet food each month.

(d) What is the chance of the survey result for the average number of people per household being 2.4 or lower if the city-wide average is really 2.5 with a standard deviation of 1.0?

7. *(10 points)* Relative to, "Poll watchers: The poll less traveled" by Richard Morin and Claudia Deane in *The Washington Post,* September 20, 2000: What types of error (other than the official "statistical error") are there for these political polls? What might have been meant by the comment that a likely voter model "works better closer to the election"?

Some Possibly Useful Formulas:

$$\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$EV \text{ of sum (count)} = n \cdot (AV \text{ of box}) \qquad SE \text{ of sum (count)} = \sqrt{n} \cdot (SD \text{ of box})$$

$$EV \text{ of avg } (\%) = AV \text{ of box} \qquad SE \text{ of avg } (\%) = \frac{SD \text{ of box}}{\sqrt{n}}$$

$$SD = (\text{larger} - \text{smaller})\sqrt{(\text{fraction with larger})(\text{fraction with smaller})}$$

Math 102 / Core 143 AX and BX — Solutions to Exam II

1. (a) $\frac{8}{48} = \frac{1}{6}$

   (b) $\frac{7}{47}$

   (c) $\frac{1}{6} \cdot \frac{7}{47} = \frac{7}{282}$

   (d) $\frac{1}{6} + \frac{1}{4} - \frac{1}{24} = \frac{3}{8}$

   (e) $1 - P(\text{no spades}) = 1 - (\frac{3}{4})^5 \approx 0.763$

2. $C(10,8)(\frac{1}{2})^8(\frac{1}{2})^2 + C(10,9)(\frac{1}{2})^9(\frac{1}{2}) + C(10,10)(\frac{1}{2})^{10} = [\frac{10 \cdot 9}{2 \cdot 1} + 10 + 1]/2^{10} = \frac{56}{1024} = \frac{7}{128}$.

3. (a) False: It should increase, because the SE of the count increases with the number of rolls.

   (b) True: The SE of the percentage decreases with the number of rolls.

   (c) True: Central Limit Theorem.

   (d) Ditto. Converting from sum to average is a linear change of variable, so it doesn't change standard units.

   (e) False: It gets closer to a "uniform distribution", flat from 1 to 8.

4. (a) EV of the sum is $100((1/5)3 + (4/5)(-1)) = -20$, with an SE of
   $(3 - (-1))\sqrt{(1/5)(4/5)}\sqrt{100} = 4(2/5)(10) = 16$.

   (b) EV of the sum is $100((1/10)7 + (9/10)(-1)) = -20$, with an SE of
   $(7 - (-1))\sqrt{(1/10)(9/10)}\sqrt{100} = 8(3/10)(10) = 24$.

   (c) Because the EV's are equal (and negative) and El has a larger SE, you are more likely to break even (or to lose bigger) playing El.

5. (a) 400

   (b) income levels

   (c) True: That is the process of "bootstrapping".

6. (a) The EV of the average is bootstrapped as $180/400 = 45\%$, and we would "give or take" the SE of the average, bootstrapped as $(1 - 0)\sqrt{(.45)(.55)}/\sqrt{400} \approx 2.5\%$.

   (b) We can't do this one: Household size apparently isn't normally distributed, because no household has fewer than 1 person, so it's not true that 31% of the data has a z-value less than 1. Thus, we can't use normal approximation to guess how many households have more than 2 people.

   (c) $\$54 \pm 2(\$20/\sqrt{400})$, or between $52 and $56.

   (d) The SE of the average is (bootstrapped as) $1.0/\sqrt{400} = 0.05$, so the probability of a sample average of 2.4 or lower is $P(z \leq (2.4 - 2.5)/0.05 = 2) = ((100 - 95)/2)\% = 2.5\%$.

7. The "statistical error" should include chance errors in sampling 95% (or maybe 68%) of the time, but about 5% (or 32%) of the time the chance error will exceed this, so that is one source of additional error. Others include various kinds of bias in the choice of sample (too many people with telephones, too few people in supermarkets rather than offices, etc.) or in the way the survey is constructed (order or wording of questions, etc.). In particular, a "likely voter" model sounds like it would discount the opinion of someone who says (s)he won't vote, but if the poll is taken a long time before election day, such a person has plenty of time for a change of mind.