NYU STERN

NEW YORK UNIVERSITY LEONARD N. STERN SCHOOL OF BUSINESS

IOMS Department

Statistics and Data Analysis

Professor William Greene
Phone: 212.998.0876
Office: KMC   7-90
Home page: http://people.stern.nyu.edu/wgreene
Email: wgreene@stern.nyu.edu
Course web page: http://people.stern.nyu.edu/wgreene/Statistics/Outline.htm

Assignment 2

Note:
The data sets for this homework (and for the other problem sets for this course) are all stored on the home page for this course. You can find links to all of them on the course outline, at the bottom with the links to the problem sets themselves.

Part I. Binomial Probability

1. Let r be a binomial random variable. Compute these probabilities using only a hand calculator. Compute $P_r(r)$ for each of the following situations:
  a. $n = 10$, $\theta = .2$, $r = 3$
  b. $n = 4$, $\theta = .4$, $r = 2$
  c. $n = 16$, $\theta = .7$, $r = 12$

2. A chain of motels has adopted a policy of giving a 3% discount to customers who pay in cash rather than by credit cards. Its experience is that 30% of all customers take the discount. Let $Y =$ number of discount takers among the next 20 customers.
  a. Do you think the binomial assumptions are reasonable in this situation?
  b. Assuming that the binomial probabilities apply, find the probability that exactly 5 of
   the next 20 customers take the discount.
  c. Find P(5 or fewer customers take the discount).
  d. What is the most probable number of discount takers in the next 20 customers?

3. The admissions office of a small, selective liberal-arts college will only offer admission to applicants who have a certain mix of accomplishments, including acombined SAT score of 1,400 or more. Based on past records, the head of admissionsfeels that the probability is 0.66 that an admitted applicant will come to the college. If 500 applicants are admitted, what is the probability that 340 or more will come? Notethat "340 or more" means the set of values {340, 341, 342, 343, …, 499, 500}.

4. Suppose that a full-repair warranty is offered with each new Power-Up food processor. If the probability that any individual food processor will be returned for needed warranty repairs within one year is 0.11, and if a certain store sells 83 of these,find the probabilities that
      a. at most 10 food processors will be returned for warranty repairs;
      b. at least 10 food processors will be returned for warranty repairs;
      c. exactly 10 food processors will be returned for warranty repairs;
      d. not more than 15 food processors will be returned for warranty repairs.

Part II. Poisson Probability

5. The rate of home sales at a small real estate agency is 1.3 per day. We'll assume that a Poisson phenomenon can represent these home sales.
      a. Find the probability that no homes will be sold on Monday.
      b. Find the probability that one home will be sold on Monday.
      c. Find the probability that two homes will be sold on Monday.
      d. Find the probability that more than two homes will be sold on Monday.

6. For each of the following situations, indicate whether the model should be binomial orPoisson or something else.
      a .The number of major forest fires to strike Colorado in calendar year 2006.
      b. The number of trading days in the month of October that the stock of General Electric will go up in value.
      c. The number of plays at craps, out of 50 attempted, that are winners.
      d. The number of prize coupons, out of 800 inserted into cereal boxes, that are returned to collect the prizes.
      e. The number of visitors to your web site on 25 FEB 2007.
      f. The number of dead squirrels found on one mile of highway 93, on May15. (Such schemes are actually used to estimate animal populations.)
      g. The number of expense account claims with inadequate documentation, in a sample of 10 selected from a master file of 280.
      h. The number of mattresses, out of 140 sold during the month of May, returned by the customers.
      i. The number of diamonds in a single hand at hearts. (In the game of hearts, a single hand consists of 13 cards dealt from the deck of 52.)
      j. The number of customers, out of 418 who made a purchase at Windham Supermarket, who purchased milk.

Part III. Normal Distribution

7. It is maintained that, in a quiet equity market with no news, the daily number of shares trades of EquiNimbus Corporation will be approximately normally distributed with mean 280,000 and with standard deviation 32,000. Find the probability that the number of shares traded tomorrow will be at most 325,000.

8. The quantity produced daily at the Milesite cement factory is approximately normally distributed with mean 0.82 and standard deviation 0.14. Production is independent from one day to the next. The units are in *millions* of pounds. Find the probability that the total production for the next 20 days will be between 16 and 17 million pounds. HINT: The total will be between 16 and 17 if and only if theaverage is between $16/20 = 0.80$ and $17/20 = 0.85$.

Part IV. Law of Large Numbers.

9. In Notes (slides) 10, we looked at the idea that in estimating a mean of a population, a larger sample is better than a small one. "Better" is quantified in the idea of the "standard error of the mean," which is computed as $\sigma/\sqrt{n}$ for a sample of n observations. A useful question to consider is "how much better?" Suppose I have drawn a sample of 10,000 observations on the number of minutes that arriving flights are late at airports around the world. I find that the sample mean is 24.75 and the sample standard deviation is 9.32. What is the estimate of the standard error of the mean? Now, the question. How much better would you say a sample of 100,000 observations would be?

10. We have found (and will continue to find) many uses for the empirical rule: 95% of almost any distribution will lie within two standard deviations of the mean. One of the ways we use this result is to form a plausible range of values around an estimate of the mean of a population that we can feel accounts for the uncertainty (sampling variability) of that estimator. For the results in problem 9 above, what would you report as your plausible range of values for the true average number of minutes late for flights assuming that the sample used is 10,000 flights?

Part V. Statistical Quality Control

11. In the DataStor case, the management gets into some trouble because of an error in statistical methodology made by one of the quality control engineers. What is the error? Describe the problem in a short paragraph. (Note, the authors of the case answer this question specifically in their description of the statistical investigation.)