Midterm

COR1-GB.1305 – Statistics and Data Analysis

The exam is open book and notes. You are also permitted use of a calculator. Each part of each problem is worth 5 points. There are 75 points total. There is no penalty for guessing incorrectly on a multiple choice problem. Partial credit may be awarded if you show work.

For the problems involving calculations, you must show all work to get full credit. For short-answer problems, there should not be any symbols in your final answer (p, n, μ, etc.), but you do not need to fully simplify your answer. It is ok to have quantities like $\binom{5}{2}$, $\sqrt{3.1}$, etc. in your final answers on these problems.

NYU Stern Honor Code:

I will not lie, cheat or steal to gain an academic advantage, or tolerate those who do.

Signature: _____ Date: _____

Name: _____

Short Answer

1. (5 points) Suppose that you collect a sample of 20 randomly-chosen Amazon reviews of the book *Gone Girl* to learn about opinions towards the book. In selecting your sample, you take care to ensure that every review of the book listed on Amazon has an equal chance of being selected. Describe a relevant population (1 sentence), and explain why your sample might be biased (1–2 sentences).

2. (5 points) Considered as a random variable, the number of visitors to the High Line Park on a typical summer day has expectation $10,000$ and standard deviation $1,500$.

   (a) Find the expected number of visitors to the High Line Park during July (total, not per day).
   Note: there are 31 days in July.

   (b) What is the interpretation of your answer to part (a)?

(c) Give an approximate range for number of visitors to the High Line Park during July (total, not per day) in a typical year. Take "typical" to mean 95%, and make other simplifying assumptions if necessary.

(d) What assumptions do you need for your answer in part (c) to be valid? Comment briefly on whether or not you think these assumptions are reasonable (1–2 sentences).

Multiple Choice

3. (5 points) If X is a normal random variable with mean $\mu = 0$ and standard deviation $\sigma = 1$, find the value of x such that $\mathrm{P}(X > x) = 0.2$.

   A. 0.5793

   B. 0.8416

   C. 1.2816

   D. 0.4207

   E. Not enough information to determine.

4. (5 points) Here is a list of 60 temperatures, measured in degrees Fahrenheit:

64.9	64.8	23.5	37.2	118.8	111.1	82.2	−10.4	69.7	−4.3	−9.2	57.6	101.1	67.5	96.6
42.6	33.6	112.8	49.5	33.3	72.0	19.9	20.3	69.7	103.7	62.8	32.7	59.3	100.3	67.4
52.3	28.9	69.3	−7.0	102.4	112.4	72.0	39.0	34.4	24.2	9.3	49.8	65.5	44.1	−13.4
70.4	29.4	26.1	118.4	−25.5	3.5	10.1	48.9	62.3	0.1	46.6	43.8	45.5	43.7	1.3

   The sample mean of these values is 49.3 and the sample standard deviation is 36.5. If we convert all of the temperatures from Fahrenheit to Celsius using the formula $C = \frac{5}{9}(F - 32)$, the sample mean and sample standard deviation of the Celsius temperatures will be:

   A. mean 9.6; standard deviation 20.3

   B. mean 9.6; standard deviation 2.5

   C. mean 49.3; standard deviation 36.5

   D. mean 9.6; standard deviation 36.5

   E. None of the above.

5. (10 points) Suppose that X is a random variable with the following probability distribution function (PDF):

x	5	10	15	20	25
$p(x)$	0.40	0.05	0.10	0.30	0.15

(a) Find the expectation $\mathrm{E}[X]$.

A. 15

B. 0.2

C. 13.75

D. 5

E. None of the above.

(b) Find $P(5.9 \leq X \leq 21.6)$.

A. 45%

B. 85%

C. 68%

D. 95%

E. Not enough information to determine.

6. (5 points) A multiple choice test has 11 questions, each with 4 possible answers. If you randomly guess on every question, what is the expected number of questions you get right?

   A. 3

   B. 2.75

   C. 0.25

   D. 2

   E. None of the above.

7. (5 points) You draw a sample of size $n = 64$ from a population with mean $\mu = 50$ and standard deviation $\sigma = 30$. Let $\bar{X}$ denote the sample mean. Approximately what is $P(\bar{X} < 56)$?

   A. 0.945

   B. 0.055

   C. 0.421

   D. 0.579

   E. Not enough information to determine.

8. (10 points) Your company has a mean annual profit of \$60MM with a standard deviation of \$5MM. Assume that the distribution of your annual profits is symmetric and mound-shaped.

   (a) Which of the following gives the range of your profits in "typical" years?

      A. \$50MM to \$70MM
      B. \$60MM to \$70MM
      C. \$55MM to \$65MM
      D. \$55MM to \$75MM
      E. None of the above.

   (b) One year you got an annual profit of \$82MM. How many standard deviations above the mean is this?

      A. 2.8
      B. 4.4
      C. 16.4
      D. 1.3
      E. None of the above.

9. (5 points) A class has 15 female students and 25 male students. If you randomly pick two different students from the class (sampling without replacement), what is the probability that you pick two females?

    A. 37.50%

    B. 13.46%

    C. 14.06%

    D. 75.00%

    E. None of the above.

10. (5 points) When a customer orders a pair of shoes on Zappos.com, there is a 10% chance that they will return their purchase. Suppose that five customers order shoes (one pair each). Approximately what is the probability that two or more customers return their orders?

    A. 2.0%

    B. 7.3%

    C. 8.1%

    D. 99.1%

    E. None of the above.