CSE 312: Foundations of Computing II
Review questions for final exam
March 12, 2015

These review questions were constructed by TAs. Any resemblance to real exam questions, living or dead, is purely coincidental.

1. A European city's temperature is modeled as a random variable with mean μ and standard deviation σ, measured on the Celsius scale. A day is described as "ordinary" if the temperature during that day remains within one standard deviation of the mean.
   (a) Give formulas for the mean and variance, if temperature is measured on the Fahrenheit scale. The formula for conversion is F = 32 + 1.8C.
   (b) From your formulas in part (a), give formulas for the temperature range for an ordinary day on the Fahrenheit scale.
2. You flip independently a fair coin and count the number of flips until the first tail, including that tail flip in the count. If the count is n, you receive 2^n dollars. What is the expected amount you will receive? How much would you be willing to pay at the start to play this game?
3. During each day, the probability that your computer's operating system crashes at least once is 5%, independent of every other day. You are interested in the probability of at least 45 crash-free days out of the next 50 days.
   (a) Find the probability of interest by using the normal approximation to the binomial.
   (b) Find the probability of interest by using the Poisson approximation to the binomial.
4. A line segment S of length L is split into two at a randomly chosen point according to the exponential distribution with $\lambda = 4/L$, treating one endpoint of S as the point 0 and the other endpoint as the point L.
   (a) Find the probability that the ratio of the shorter to the longer segment is less than 1/3. (If the splitting point is less than 0 or greater than L, treat that as though the ratio is less than 1/3.)
   (b) What is the probability that the splitting point is less than 0 or greater than L?
5. A computer network consisting of n computers is to be formed by connecting each computer to each of the others by a direct ("point-to-point") network cable.
   (a) How many network cables are needed?

(b) Unfortunately, some of the cables may be faulty ("dead") while others are OK ("alive"). How many different "connectivity patterns" are possible? (E.g., "the cable between computers 1 and 3 is alive, but no others are" is one pattern; "between 1 and 4, but no others" is a different pattern; "only the cable between 1 and 4 is dead" is a third pattern, etc.)

(c) Assuming that there is at least one "live" cable connected to every computer, show that there are at least two computers in the network that are directly connected to the same number of other computers via live cables.

6. Alice, Bob, and Carol repeatedly take turns rolling a fair die. Alice begins, Bob always follows Alice, Carol always follows Bob, and Alice always follows Carol. Find the probability that Carol will be the first one to roll a six.

7. A line segment S of length L is split into two at a randomly chosen point according to the normal distribution with mean $\mu = L/2$ and standard deviation $\sigma = L/4$, treating one endpoint of S as the point 0 and the other endpoint as the point L.

   (a) Find the probability that the ratio of the shorter to the longer segment is less than 1/3. (If the splitting point is less than 0 or greater than L, treat that as though the ratio is less than 1/3.)

   (b) What is the probability that the splitting point is less than 0 or greater than L?

8. (a) Suppose $x_1, x_2, \ldots, x_n$ are samples from a normal distribution whose mean is known to be zero, but whose variance is unknown. What is the maximum likelihood estimator for its variance?

   (b) Suppose the mean is known to be μ but the variance is unknown. How does the maximum likelihood estimator for the variance differ from the maximum likelihood estimator when both mean and variance are unknown?

9. Let $f(x \mid \theta) = \theta x^{\theta-1}$ for $0 \leq x \leq 1$, where θ is any positive real number. Let $x_1, x_2, \ldots, x_n$ be i.i.d. samples from this distribution. Derive the maximum likelihood estimator $\hat{\theta}$.

10. The number of seconds a server takes to finish a job is modeled as a random variable X from an unknown distribution. You would like to be able to guarantee clients that, with high probability, jobs will be finished within 50 seconds. What is the best guarantee you could give if:

    (a) You assume that X has mean 25.

    (b) You assume that X has mean 25 and variance 25.

    (c) You assume that $X \sim \text{Poi}(25)$. (Hint: use the Normal approximation of the Poisson. Why is it reasonable to approximate Poi(25) by a normal distribution? It follows from the central limit theorem, since it turns out that a Poisson random variable with $\lambda = 25$ is the sum of 25 independent Poisson random variables each with $\lambda = 1$. See `https://onlinecourses.science.psu.edu/stat414/node/180`.

11. A frog starts at position 0 on a line and at each second t jumps X_t cm, where the X_t are all i.i.d. according to the following probability mass function:

$$\begin{aligned} p(-2) &= 1/6 \\ p(-1) &= 1/3 \\ p(1) &= 1/6 \\ p(2) &= 1/3 \end{aligned}$$

Use the central limit theorem to estimate the probability that, after 100 jumps, the frog is at a negative position.

12. Chebyshev's inequality implies that the proportion of observations that are at most 3 standard deviations from the mean is at least p. Determine the value of p.

13. Related to HW6, problem 3: You throw a dart at a circular target of radius $r = 5$ inches. Your aim is such that the dart is equally likely to hit any point in the target. For each throw, you win \$1 if the dart strikes within 2 inches of the target's center. Let W be your total winnings for 100 independent throws. Use the Chernoff bound to get an upper bound on the probability that you win at least \$24. (The Chernoff bound was given in the form $\mathrm{P}(X > (1+\delta)\mu) \leq \dots$, but the same bound actually also holds in the form $\mathrm{P}(X \geq (1+\delta)\mu) \leq \dots$.)