STAT2800/2010: A Review of Some Distributions

You have studied these common probability distributions in your course:

Binomial Distribution

- n independent trials, proportion of success p, two outcomes (success, failure).

$$p(x) = P(X = x) = \begin{cases} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, x = 0,1,\ldots,n \\ 0, \quad \text{otherwise} \end{cases}$$

Poisson Distribution

- λ is the average number of occurrences in the specified period, an approximation of the binomial when n is large, and p is small. Example: Find the probability of 5 people waiting in a queue at the bank on a weekday if the average is 3.

$$p(x) = P(X = x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, \text{ for } x = 0,\ 1,\ 2,\ \ldots \\ 0, \quad \text{otherwise} \end{cases}$$

Exponential Distribution

- Most common distribution for wait times (how long will the duration be?). Example: If the average wait time is 5 minutes, find the probability of spending more than 10 minutes in queue.
- $p(x > c) = e^{-\lambda c}$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, \ x > 0 \\ 0, \quad \text{otherwise} \end{cases}.$$

Normal Distribution

- Many natural and social occurrences follow this distribution.
- We convert to standard units, z – scores, to use a table.

Lognormal Distribution

- Not symmetrical, if Y has a lognormal distribution, then X = ln(Y) has a normal distribution.

Weibull Distribution

- Commonly models the lifetime of components. When α=1, this is the exponential distribution.
- $p(x < t) = 1 - e^{-(\beta t)^{\alpha}}$

$$f(x) = \begin{cases} \alpha\beta^{\alpha} x^{\alpha-1} e^{-(\beta x)^{\alpha}}, x > 0 \\ 0, \quad x \le 0 \end{cases}.$$

1. Of all the registered automobiles in a certain province, 20% violate the province emissions standard. Twelve automobiles are selected at random to undergo an emissions test. Find the probability that fewer than three of them fail test. (Ans: 0.5583)

2. The time between requests to a web server is exponentially distributed with mean 0.4 seconds. What is the 80th percentile? (Ans: 0.6438 s)

3. If X is a normally distributed variable with $\mu = 24$ and $\sigma = 2.3$, find the following:
   a) $P(X < 20)$ [Ans: 0.0409]

   b) $P(X > 27)$ [Ans: 0.0968]

   c) $P(19 < X < 28)$

Select the correct answer:

i) 0.0150
ii) 0.9441
iii) 0.9599
iv) 0.9850

Select all that are true about the normal distribution:

a) For any normal distribution, the mean, median, and mode will have the same value
b) In a normal distribution, about 60% of scores are greater than Z = 0.25
c) The normal distribution curve can be used to determine probability only for normally distributed populations
d) The percentile rank for the mean is 50% for any normal distribution
e) A Z-score represents the number of standard deviations above or below the mean.

4. The distance between consecutive flaws on a roll of sheet aluminum is exponentially distributed with mean distance 3 m. What is the probability that a 5 m length of aluminum contains exactly two flaws? (Ans: 0.2623)

5. Of the bolts manufactured for a certain application, 81% meet the length specification and can be used immediately, 5% are too long and can be used after being cut, and 14% are too short and must be scrapped. Find the probability that fewer than 9 out of a sample of 10 bolts can be used (either immediately or after being cut). (Ans: 0.4185)

6. An article models the increase in the risk of cancer due to exposure to carbon tetrachloride as lognormal with μ = -14.45 and σ = 0.79. Find the 95th percentile. (Ans: 1.945×10^{-6})

7. The lifetime, in hours, of a certain type of bearing is modeled with the Weibull distribution with parameters α= 2.35 and β= 4.474×10^{-4}.
   a) Find the probability that a bearing lasts more than 1000 hours. (Ans: 0.8598)

   b) Find the median lifetime of a bearing. (Ans: 1905.9 hrs)

8. The number of typos on the webpage of the Toronto Star has a Poisson distribution with a mean of 1.2 errors per page.
   i) Select the probability that there are less than two errors on the first page:
      a) 0.2169
      b) 0.3011
      c) 0.3614
      d) 0.6626

   ii) Find the probability that is there at least one error in total on the first two webpages viewed. (Ans:0.9093)

   iii) Find the probability that two successive webpages each contain two errors. (Ans: 0.0470)

9. For an annularly threaded nail driven into spruce-pine-fir lumber, the ultimate removal strength (N/mm) was modeled as lognormal with $\mu = 3.8$ and $\sigma = 0.219$. For a helically threaded nail under the same conditions, the strength was modelled as lognormal with $\mu = 3.37$ and $\sigma = 0.272$. What is the probability that a helically threaded nail will have a **greater** removal strength than the **median** for annularly threaded nails? (Ans: 0.0571)

10. A catalyst researcher states that the diameters, in microns, of the pores in a new product she has made have the exponential distribution with parameter $\lambda = 0.25$. What is the median pore diameter? (Ans: 2.7726 microns)

11. A system consists of two components connected in series. The system will fail when either of the two components fails. Let T be the time at which the system fails. Let X_1 and X_2 be the lifetimes of the two components. Assume that X_1 and X_2 are independent and that each has the Weibull distribution with $\alpha = 2$ and $\beta = 0.25$.

    Find $P(X_1 > 5 \text{ and } X_2 > 5)$. (Ans: 0.04394)

12. A new process has been designed to make ceramic tiles. The goal is to have no more than 5% of the tiles be nonconforming due to surface defects. A random sample of 1000 tiles is inspected. Let X be the number of nonconforming tiles in the sample. If 5% of the tiles produced are nonconforming, what is $P(X \geq 75)$? (Ans: 0.0002)