Practice Exam 2 (Posted on Mar. 20, 2015) **Name:**

1. The time between arrivals of taxis at a busy intersection is exponentially distributed with a **mean** of 10 minutes.

(a) What is the probability that you wait longer than half hour for a taxi?

(b) Suppose that you have already been waiting for half hour for a taxi. What is the probability that one arrives within the next 5 minutes?

(c) Determine x such that the probability that you wait more than x minutes is 0.95?

Sol. Here we define a random variable T to be the time between arrivals of taxis at a busy intersection, then

$$T \sim \text{Exp}(\lambda).$$

Since $E(T) = 1/\lambda = 10$, then we have $\lambda = 0.1$, i.e., $T \sim \text{Exp}\ (\lambda = 0.1)$.
(a) $P(X > 30) = e^{-0.1(30)} = 0.0498$.

(b) Since exponential distributions have memoryless property, $P(X > 30+5|X > 30) = P(X > 5) = e^{-0.1(5)} = e^{-0.5} = 0.6065$. Then$P(X < 30 + 5|X > 30) = 1 - P(X > 5) = 1 - e^{-0.1(5)} = 1 - e^{-0.5} = 1 - 0.6065 = 0.3935$.

(c) $0.95 = P(T > x) = e^{-0.1x}$ and we solve $x = -10 \times \log 0.95 = 0.5129$.

2. A typesetter, on the average, makes one error in every 500 words typeset. A typical page contains 300 words. Define Y as the number of errors in a page which follows a Poisson distribution with a mean 0.6, i.e.,

$$Y \sim \text{Poisson}(0.6)$$

Determine the following:

(a) Mean pages until the first error.

(b) Probability that we will have to wait more than 15 pages to see the first error?

(c) Mean pages until the 30th error.

(d) Mean pages between error 30 and 45.

Sol.
(a) Define W_1 be the number of pages to see the first error, i.e.,

$$W_1 \sim \text{Exp}(\lambda = 3/5).$$

Then we have

$$E(W_1) = 1/\lambda = 5/3.$$

(b)

$$P(W_1 > 15) = \exp\{-(3/5) \times 15\} = \exp\{-9\} = 1.2341 \times 10^{-4}.$$

(c) Define W_2 be the number of pages to see the 30th error, i.e.,

$$W_2 \sim \text{Gamma}(\alpha = 30, \lambda = 3/5).$$

Then we have

$$E(W_2) = \alpha/\lambda = 30/(3/5) = 50.$$

(d) Define W_3 be the number of pages to see the first error, i.e.,

$$W_3 \sim \text{Gamma}(\alpha = 45 - 30, \lambda = 3/5).$$

Then we have

$$E(W_3) = \alpha/\lambda = 15/(3/5) = 25.$$

3. A driver's reaction time to visual stimulus is normally distributed with a mean of 0.4 seconds and a standard deviation of 0.05 seconds.

(a) What is the probability that a reaction requires more than 0.5 seconds?

(b) What is the probability that a reaction requires between 0.5 and 0.6 seconds?

(c) What reaction time is exceeded 95% of the time?

Sol. Define X to be a driver's reaction time to visual stimulus and $X \sim N(\mu = 0.4, \sigma^2 = (0.05)^2)$.

(a) $P(X > 0.5) = \texttt{normalcdf}(0.5, 10^{99}, 0.4, 0.05) = 0.0228.$

(b) $P(0.5 < X < 0.6) = \texttt{normalcdf}(0.5, 0.6, 0.4, 0.05) = 0.0227.$

(c) Find x such that $0.95 = P(X < x)$ and $x = \texttt{invNorm}(0.95, 0.4, 0.05) = 0.4822.$

4. Define

$$X = \text{the current measurements in a strip of wire.}$$

We assume

$$X \sim N(\mu = 3, \sigma^2 = 0.25).$$

Suppose that we take a random sample of $n = 16$ in a strip of wire $X_1, \ldots, X_{16}$.

(a) What is the distribution of the sample mean $\overline{X}$?

(b) Is $\overline{X}$ an unbiased estimator of μ?

(c) Find $P(2.75 < \overline{X} < 3.25)$.

Sol.

(a) The sampling distribution of $\overline{X}$ is

$$\overline{X} \sim N\left(\mu = 3, \left(\frac{\sigma}{\sqrt{n}}\right)^2 = \left(\frac{0.5}{\sqrt{16}}\right)^2\right).$$

That is, we have $\overline{X} \sim N(3, 0.125)$.

(b) According to (a), since $E(\overline{X}) = \mu$, $\overline{X}$ is an unbiased estimator of μ.

(c) According to (a), $P(2.75 < \overline{X} < 3.25) = \texttt{normalcdf}(2.75, 3.25, 3, 0.125) = 0.5662$.

5. The amount of time that a customer spends waiting at an airport check-in counter is a random variable X with mean 8.2 minutes and standard deviation 1.5 minutes. Suppose that a random sample $X_1, X_2, \ldots, X_{49}$ of $n = 49$ customers is observed.

(a) Please describe the central limit theorem (CLT).

(b) Find the probability that the average time waiting in line $\overline{X}$ is less than 5 minutes.

(c) Find the probability that the average time waiting in line $\overline{X}$ is between 5 and 10 minutes.

(d) Find the probability that the average time waiting in line $\overline{X}$ is greater than 10 minutes.

Sol.

(a) $X_1, \ldots, X_n$ is a random sample from a population with mean μ and variance σ^2. The sampling distribution of $\overline{X}$ is approximated normal, i.e.,

$$\overline{X} \sim AN\left(\mu, \frac{\sigma^2}{n}\right).$$

(b) According to CLT, we have $\overline{X} \sim AN(\mu = 8.2, \sigma^2 = 1.5^2/49)$, i.e, $\overline{X} \sim AN(8.2, 0.2143)$. So $P(X < 5) = \texttt{normcdf}(-10^{99}, 5, 8.2, 0.2143) \approx 0$.

(c) Similar to (a), $P(5 < X < 10) = \texttt{normcdf}(5, 10, 8.2, 0.2143) \approx 1$.

(d) Similar to (a), $P(10 < X) = \texttt{normcdf}(10, 10^{99}, 8.2, 0.2143) \approx 0$.

6. The life (in hour) of a magnetic resonance imaging machine (MRI) is modeled by a Weibull distribution with shape and scale parameters β and η. The following data are $n = 30$ time to failure for MRI:

203	225	237	283	287	293	341	347	349	350
362	388	400	425	426	453	457	484	504	550
553	560	576	584	622	636	658	730	797	826

Under a Weibull assumption for

$$T = \text{time (in hour) to failure of MRI},$$

we calculate the maximum likelihood estimate $\hat{\beta} = 3.06$ and $\hat{\eta} = 518.86$ for the data above. Use these values (along with the Weibull assumption) to answer the following questions.

(a) Calculate $P(T > 250)$. Interpret what is this probability means in words.

(b) According to the estimated hazard function for T (Figure 1). Explain, in plain English, what information this graph reveals.

Sol.

(a)

$$P(T > 250) = e^{-(250/518.86)^{3.06}} = 0.8985.$$

It is the probability a MRI machine will survive past 250 hours.

(b) Since $\hat{\beta} = 3.06 > 1$ and the hazard function graph in Figure 1 , we see that the hazard function is increasing. So the population of MRI machine should be getting weaker overtime. Furthermore, the rate of failure increases at an increasing rate.

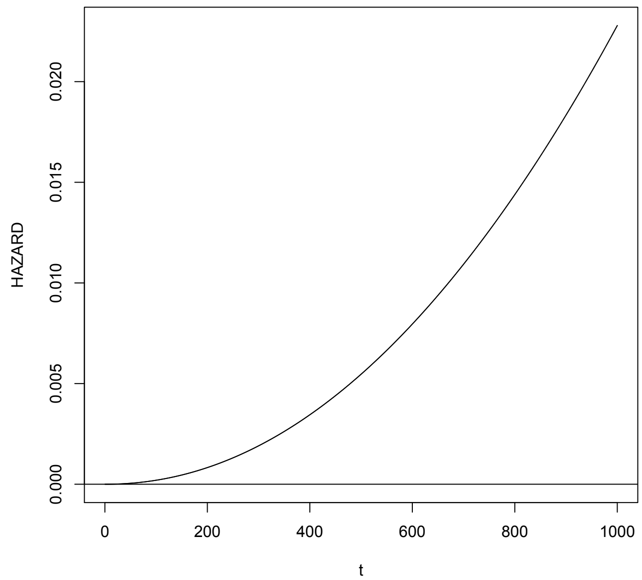

Figure 1: Hazard function for MRI data.

7. Data on the concentration of polychlorinated biphenyl (PCB) residues in a series of lake trout from Cayuga Lake, NY, can be found in Bates and Watts (1988). Each whole fish was mechanically chopped, ground, and thoroughly mixed, and 5–gram samples taken. The samples were treated and PCB residues in parts per million (ppm) were estimated using column chromatography. Historically, the PCB residues have a standard deviation of 0.7 ppm. Assume that the PCB residues are normally distributed. A subset of the random sample of the PCB residues was measured in ppm: 0.6, 1.6, 0.5, 1.2, 2.0, 1.3, 2.5, 2.2, 2.4, 1.2. An important question of interest is whether the population mean of PCB residues exceed 1.4.

(a) What are the sample mean $\bar{x}$, sample variance s^2, and sample standard deviation s?

(b) Construct a 95% two-sided confidence intervals for the population mean of PCB residues μ and give an interpretation.

(c) Construct a 95% one-sided lower-confidence bound for the population mean PCB residues μ and give an interpretation.

Bonus! According to the lower-confidence bound in (c), could we say that μ exceed 1.4? Why?
Sol.
(a) $\overline{x} = 1.55$, $s^2 = 0.5072$, and $s = 0.7122$.

(b) A 95% two-sided confidence interval: $(1.116, 1.984)$. We are 95% confident that the population mean of PCB residues are between 1.116 and 1.984.

(c) A 95% one-sided lower-confidence bound: $(1.186, \infty)$. We are 95% confident that the population mean of PCB residues are larger than 1.186.