Exercise set 3

Stat 251

Instructions: Use StatDistributions.com for probabilities

(1) Find the following normal probabilities (use StatDistributions.com)
  (a) $P(Z < -1.05)$
  (b) $P(Z > 2)$
  (c) $P(Z > 8)$
  (d) $P(0 \leq Z \leq 2.17)$
  (e) Find the z score that represents the top 5%
  (f) Find the z score that represents $Q1$

(2) Suppose the force acting upon a column that helps to support a building is normally distributed with mean 15.0 kips and standard deviation 1.25 kips (i.e. $X \sim N(\mu, \sigma) = X \sim N(15, 1.25)$). Calculate the following:
  (a) $P(X \leq 15)$
  (b) $P(X \leq 17)$
  (c) $P(X \geq 10)$
  (d) $P(14 < X < 18)$
  (e) Find the value of the force (in kips) that the highest 7% of forces on the column are

(3) Salaries for teachers in a particular elementary school district are normally distributed with a mean of $44,000 and a standard deviation of $6,500. We randomly survey ten teachers from that district.
  (a) In words, state what X is
  (b) Give the distribution of X in shorthand notation
  (c) In words, state what the total ($\hat{\tau}$) is
  (d) Give the distribution of $\hat{\tau}$ in shorthand notation
  (e) Find the probability that the teachers earn a *total* of over $400,000.
  (f) Find the 90th percentile for an *individual* teacher's salary
  (g) Find the 90th percentile for the *sum* of ten teachers' salary
  (h) If we surveyed 70 teachers instead of ten, graphically, how would that change the distribution in part d?
  (i) If each of the 70 teachers received a $3,000 raise, graphically, how would that change the distribution in part b?

(4) The average length of a maternity stay in a U.S. hospital is said to be 2.4 days with a standard deviation of 0.9 days. We randomly survey 80 women who recently bore children in a U.S. hospital.
  (a) In words, state what X is
  (b) Give the distribution of X in shorthand notation
  (c) In words, state what $\overline{X}$ is
  (d) Give the distribution of $\overline{X}$ in shorthand notation
  (e) In words, state what $\hat{\tau}$ is
  (f) Give the distribution of $\hat{\tau}$ in shorthand notation
  (g) Is it likely that an *individual* stayed more than five days in the hospital? Why or why not? Justify with calculations.
  (h) Is it likely that the *average* stay for the 80 women was more than five days? Why or why not? Justify with calculations.
  (i) Which is more likely (justify with calculations):
    (1) An *individual* stayed more than five days

(2) the *average* stay of 80 women was more than five days

(j) If we were to sum up the women's stays, is it likely that, collectively they spent more than a year in hospital? Why or why not? Justify with calculations.

(5) *Finding Dory*

Coral Reef communities are home to one-quarter of all marine plants and animals worldwide. These reef support large fisheries by providing breeding grounds and safe havens for young fish of many species. Coral reefs are seawalls (protecting shorelines from tides, storm-surges, and hurricanes as well as produce the limestone and sand of which beaches are made), Marine scientists say that a tenth of the world's reef systems have been destroyed in recent times. At current rates of loss, almost three-quarters of the reefs could be gone in 30 years.

A particular lab studies corals and the diseases that affect them. Dr. Drew Harvell and his lab sampled sea fans at 19 randomly selected reefs along the Yucatan peninsula and diagnosed whether the animals (the sea fans) were affected by *aspergillosis*[1]. In specimens collected at a depth of 40 feet at the Las Redes Reef in Akumal, Mexico, scientists found that 52% of the 104 sampled sea fans were infected with *aspergillosis*.

(a) What are the mean (proportion, π) and standard deviation of the sampling distribution of the sample proportion (mean (π) and $se_{\hat{\pi}}$) of infected sea fans? What should the distribution look like (think of the definition of CLT)?

(b) What is probability that the sample proportion of infected sea fans is less than 50% (that is find $P(\hat{\pi} < .5)$)?

(c) What is probability that the sample proportion of infected sea fans is between 40 and 60%?

(d) What is the proportion of infected sea fans that represents the 90^{th} percentile?

(6) *There is no Dana, only Zeul (Who you gonna call?)*

In November of 2005 the Harris Poll asked 889 randomly selected US adults, "Do you believe in ghosts?" 29% said they did.

(a) In constructing confidence intervals, would we use $z^\star$ or $t^\star$ in this situation? Briefly explain why you would use one instead of the other.

(b) Estimate π, the true proportion of US adults that believe in ghosts, with 90% confidence. Interpret the interval in context of the data.

(c) Suppose, using the information from the survey (the 29% that believe in ghosts) that a new survey is to be taken and the new bound is to be 2%. What sample size will be required?

(d) Suppose that we know nothing of any prior results from this study (thus have no estimate for the proportion of those US adults that believe in ghosts). What proportion should we use for the estimation? What sample size do we need with no prior information? Why is it different than the sample size from part (c)?

[1] K.M.Mullen, C.D.Harvell, A.P.Alker, D. Dube, E. Jordán-Dahlgren, J. R. Ward, and L.E. Petes, "Host range and resistance to aspergillosis in three sea fan species from the Yucatan", *Marine Biology* (2006), Springer-Verlag.