Exercise 4 – MICRO 110 Spring 2024

This assignment may be completed without Python. The word "Python" will appear on any question that you may prefer to use the provided Jupyter Notebook.

This assignment may be completed without Python. The word "Python" will appear on any question that you may prefer to use the provided Jupyter Notebook.

For these problems, you may find it useful to use an online calculator such as:

http://onlinestatbook.com/2/calculators/normal_dist.html

https://stattrek.com/online-calculator/normal.aspx

1) *Python*. If a random variable has a normal distribution with mean = 80 and standard deviation = 20, what is the probability that it assumes the following values?

a) Less than 77 .4

0.45

b) Between 61.4 and 72.9

0.185

c) Greater than 90.0

0.309

d) Less than 67.6 or greater than 88.8

0.598

e) Between 92.1 and 95.4

0.052

f) Between 75.0 and 84.2

.182

g) Exactly 81.7

0

2) Suppose that 95% of the bags of a certain fertilizer mix weigh between 49 and 53 kg. Averages of three successive bags were plotted, and 47.5% of these were observed to lie between 51 and X kg. Estimate the

value of X. State the assumptions you make and say whether these assumptions are likely to be true for this example.

Assuming a normal distribution, the mean weight should be 51kg and the standard deviation should be approximately 1kg (since the 95% confidence interval is at ±2σ).

Since we are using successive averaging we need to use the statistical tools pertaining to sample distributions (page 82, Module 2). If we assume truly random averaging, the mean and standard deviation for the averaged samples should be 51kg (being as the sample mean should be the parent distribution's mean) and 1/√3 (sample distribution's standard deviation n = σ / sqrt(n)).

Based on this normal distribution, the interval spanning 47.5% is in other words +2σ will be from 51kg to approximately 52.1kg, so X = 52.1.

For -2σ, X would have been 49.9.

3) *Python.* The lengths of bolts produced in a factory may be taken to be normally distributed. The bolts are checked on two "go, no-go" gauges, where bolts that are shorter than 2.9cm or longer than 3.1cm are rejected.

a) A random sample of 397 bolts are checked on the gauges. If the mean length of the bolts produced at that time was 3.06 cm and the standard deviation was 0.03 cm, what values would you expect for n1, the number of bolts found to be too short, and n2, the number of bolts found to be too long?

n_1 will be 0, since it is approx. -5σ below the mean n_2 will be 36, based on the normal distribution.

b) A random sample of 50 bolts from another factory are also checked. If for these n1 = 12 and n2 = 12. Estimate the mean and the standard deviation for these bolts. State your assumptions.

If we assume truly random sampling, and assume a normal distribution, the mean will be exactly in the middle of 2.9 and 3.1, since the normal distribution is symmetric. Therefore, η=3cm. We can then calculate the standard deviation by noting that p(<2.9)=12/50=0.24 (this is not the z-score). Therefore, σ = 0.14 (page 69, module 1).

4) Four samples are taken and averaged every hour from a production line and the hourly measurement of a particular impurity is recorded. Approximately one out of six of these averages exceeds 1.5% impurity concentration, with the mean value approximately 1.4%. State assumptions that would enable you to determine the proportion of the individual readings exceeding 1.6% concentration. Make the assumptions and do the calculations. Are these assumptions likely to be true? If not, how might they be violated?

If we assume a normal distribution and random sampling, we can use the equation $s = \sigma/\sqrt{n}$ where n=4. The mean should be 1.4% for both the sample average distribution and the parent distribution. Using an online calculator, we can find the standard deviation of the sample average distribution that results in p(x>1.5)=1/6. This is found to be approximately 0.1. Then, the standard deviation for the parent distribution should be 0.2. For this value of standard deviation, the proportion of individual readings exceeding 1.6% is approximately 0.16.

Since the samples are taken hourly, the most likely violation is that of random sampling, since it is likely that the production line while have a time-dependent behavior. Also, if there is a dominant impurity introduction process, it is possible that the impurity distribution itself is non-normal, though this is typically unlikely.