ENGI 4421 Probability and Statistics
Faculty of Engineering and Applied Science

Problem Set 8

Confidence Intervals

---

1. In a random sample of 16 cables drawn from a large population of old cables, 12 of the cables pass a new, more stringent, strength test. Construct a 95% confidence interval for the proportion p of cables in the population that would pass the new strength test.

---

2. Provided the expected numbers of successes and failures both exceed 10, the normal distribution with parameters $\mu = np$ and $\sigma^2 = npq$ is an acceptable approximation to the binomial distribution with parameters (n, p).

   Use the normal approximation to find $\mathrm{P}\left[15 \le X < 20\right]$ when $\mathrm{P}\left[X = x\right] = b\left(x; 40, .375\right)$ **and** find $\mathrm{P}\left[15 \le X < 20\right]$ using the binomial distribution.

---

3. Measurements of the power output of a motor are known to have a standard deviation of $\sigma = 45\ \mathrm{W}$. Find the smallest sample size needed to ensure that the sample mean of a random sample lies within 10 W of the unknown true mean μ at least 95% of the time.

---

4. Prior experience leads an investigator to believe that the breaking strengths of a particular type of fibre are normally distributed, with a mean of 100 N. The strength of that belief is represented by a standard deviation of 5 N.
   A random sample of five fibres is tested to destruction. The population variance is known to be $\sigma^2 = 25\ \mathrm{N}^2$. The observed sample mean is 110**.**0 N.

   (a) Find the 95% classical confidence interval for the population mean breaking strength of the fibres.
   (b) Find the 95% Bayesian confidence interval for the population mean breaking strength of the fibres.
   (c) Comment briefly on any difference between these two confidence intervals.

---

5. An engineer is studying the fatigue life, Y, in cycles of stress, of a steel connector device. It is convenient to consider the quantity $X = \log Y$, which should be considered in the following (*i.e.* use the quantity X in the problem). Previous measurements have shown that X is a normally distributed process with a standard deviation of 0.20. The mean of X is the subject of the study, and this is denoted as "A". A first set of measurements results in the engineer assigning a normal distribution to A with a mean of 5.80, with the strength of that belief being represented by a standard deviation of 0.10.

   The following ten measurements were then taken.

   5.61, 5.24, 5.12, 5.40, 5.14, 5.38, 5.40, 5.46, 5.41, 5.67.

   (a) Determine the posterior distribution of A.
   (b) Sketch (or plot) prior and posterior probability distributions of A.
   (c) What is the name for the distribution of Y?

---

6. The mass of a prototype component is known to be normally distributed. A random sample of 100 components has a sample mean of 127.5 g and a standard deviation of 2.3 g .

   Find a 99% [classical] confidence interval estimate for the true mean mass of the prototype component.

---

7. The lifetime of an expensive filament is known to be normally distributed, to an excellent approximation, but no other prior information is available. A random sample of 5 filaments has a sample mean lifetime of 968 hours and a standard deviation of 27 hours.

   (a) Find a 95% confidence interval estimate for the true mean lifetime of the filament.
   (b) Is this sample consistent with a population mean lifetime of 1000 hours?

---

8. [Navidi, exercises 5.6, question 7]

   During the spring of 1999, many fuel storage facilities in Serbia were destroyed by bombing. As a result, significant quantities of oil products were spilled and burned, resulting in soil pollution. The article "Mobility of Heavy Metals Originating from Bombing of Industrial Sites" (B. Škrbić, J. Novaković and N. Miljević, Journal of Environmental Science and Health, 2002:7-16) reports measurements of heavy metal concentrations at several industrial sites in June 1999, just after the bombing, and again in March of 2000. At the Smederevo site, on the banks of the Danube river, eight soil specimens taken in 1999 had an average lead concentration (in mg/kg) of 10.7 with a standard deviation of 3.3. Four specimens taken in 2000 had an average lead concentration of 33.8 with a standard deviation of 0.50. Find a 95% confidence interval for the increase in lead concentration between June 1999 and March 2000.

---

9. [*Bonus question*]
A random sample of 98 runs of an existing chemical process produced 61 runs that meet a more stringent standard for its product. A random sample of 18 runs of a prototype new process had 15 runs that meet the new standard.

(a) Construct a one-sided 95% confidence interval $\left(c < p_{\text{new}} - p_{\text{old}} < 1\right)$ for the difference in population proportions of acceptable runs between the new and old processes.

(b) Can you conclude, at a 95% level of confidence, that the new process meets the standard more frequently than the old process does?

---

10. The fuel used by a random sample of seven boilers was measured first when they were operated for ten minutes from a cold start, then when they were operated for ten minutes at their steady-state [hot] temperature. The results (in cm^3) are listed below. You may assume that both populations are normally distributed.

Cold	**Hot**
308.736	294.877
208.133	191.730
223.099	208.827
238.190	205.523
313.124	320.122
379.309	346.855
290.134	219.225

At a 95% level of confidence, is there sufficient evidence to conclude that there is a difference in the true mean amount of fuel used by these boilers between hot and cold starts? Justify your choice of test.

☜ Back to the index of questions On to the solutions to this problem set ☞