MAST90105: Lab and Workshop Problems for Week 10

The Lab and Workshop this week covers problems arising from Module 7. The problems will be assigned to groups this week.

1 Lab

Last week's lab had a lot of Mathematica and R problems. So this week only 2 have been set.

1. How good are confidence intervals? If we repeat the experiment a large number of times we expect 95% of the confidence intervals for contain the parameter values. We can check this using simulations. Enter the following commands:

```
x = t.test(rnorm(10))
x
names(x)
x$conf.int
```

   You should use the help function or your tutor to understand what each command does. Note that rnorm simulates values from $N(0,1)$ so we know the true mean is zero. Then automate the process

```
 f=function(t){x=t.test(rnorm(t));as.vector(x$conf.int)};
f(10);
 f(20);
t <- as.matrix(rep(10,100));
C <- t(apply(t,1,f)); #this is a trick so we don't have to program
 matplot(C,type="l");#a matrix plot
abline(0,0)#includes a line at 0
```

   Each column of the matrix C is the lower and upper bounds of a 95% confidence interval. From your plot determine how many of these intervals contain the true mean zero. Is it close to 95%? You can check as follows:

```
num = (C[, 1] < 0) & (C[, 2] > 0)
sum(num)/nrow(C)
```

2. In class, we discussed the Newspoll outcomes from March 20 and April 3 2017. The March 20 poll reported that 675 of 1824 voters would vote first for the Government if an election were held then, and on April 3 it was 615 out of 1708 voters.

   a. Starting with a uniform distribution over (0,1), find the posterior distribution for the population proportion after the March 20 .

b. Use this posterior as a prior distribution for the April 3 Newspoll and find the resulting posterior distribution.

c. Plot this density with the posterior density obtained in lectures from a uniform prior.

d. Find a 95% posterior probability interval from your posterior distribution and compare this to the one from lectures.

e. Construct a Beta distribution as a prior for the data that arrived on April 3 based on your Bayes estimates from the previous poll so that there is 99% probablity that the true proportion is less than (a) 50% (b) 40%. Compute the posterior in each case.

2 Workshop

3. Let $X_1, \ldots, X_n$ be a random sample from a gamma distribution with $\alpha = 4$ so that

$$f(x; \theta) = \frac{1}{6\theta^4} x^3 e^{-x/\theta}, \ 0 < x < \infty, \ 0 < \theta$$

. Continuing the question from last week,give an approximate $100(1-\alpha)\%$ confidence interval for θ. *Use Mathematica to compute the appropriate derivatives and means.*

4. A random sample of size 16 from $N(\mu, 25)$ yielded $\bar{x} = 73.8$. Find a 95% confidence interval for μ. (Recall $z_{0.025} = 1.96$, $z_{0.05} = 1.645$).

5. A pet store sells guinea pig food in "2-pound" bags that are weighed on a an old 25-pound scale. Suppose it is known that the standard deviation of weights is $\sigma = 0.12$ pound. If a sample of 16 bags of guinea pig food were carefully weighed in a laboratory and the average weight was $\bar{x} = 2.09$ pounds, find an approximate 95% confidence interval for μ, the mean weight of gerbil food in the "2-pound" bags sold by the pet store.

6. To determine whether bacteria count was lower in the west basin of Lake Macatawa than in the east basin, $n = 37$ samples of water were taken in the west basin, and the number of bacteria colonies in 100 millilitres of water was counted. The sample characteristics were $\bar{x} = 11.95$ and $s = 11.80$, measured in hundreds of colonies. Find the approximate 95% confidence interval for the mean number of colonies, say μ_W, in 100 millilitres of water in the west basin. (Note, $t_{0.025}(36) = 2.028$, $t_{0.05}(36) = 1.688$)

7. Thirteen tons of cheese is stored in some old gypsum mines, including "22-pound" wheels (label weight). A random sample of $n = 9$ of these wheels yields $\bar{x} = 20.9$ and $s = 1.858$. Assuming that the weights of the wheels is $N(\mu, \sigma^2)$ find a 95% confidence interval for μ. Is the claim these are "22 pound" wheels reasonable? ($t_{0.025}(8) = 2.306$, $t_{0.05}(8) = 1.859$)

8. The length of life of brand X light bulbs is assumed to be $N(\mu_X, 784)$. The length of life of brand Y light bulbs is assumed to be $N(\mu_Y, 627)$ and these lifetimes are independent of X. If a random sample of $n = 56$ brand X light bulbs yielded $\bar{x} = 937.4$ hours and a random sample of size $m = 57$ brand Y light bulbs yielded $\bar{y} = 988.9$, find a 95% confidence interval for $\mu_X - \mu_Y$. Is it reasonable to conclude that the two brands of light bulb have the same mean lifetimes?

9. A test was conducted to determine if a wedge on the end of a plug designed to hold a seal onto that plug was operating correctly. The data were the force required to remove a seal from the plug with the wedge in place (X) and without the wedge (Y). Assume the distributions of X and Y are $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$ respectively. Samples of size 10 on each variable yielded:

Variable	n	$\bar{x}$	s
X	10	2.548	0.323
Y	10	1.564	0.210

   a. Find a 95% confidence interval for $\mu_X - \mu_Y$. ($t_{0.025}(18) = 2.101$, $t_{0.05}(18) = 1.734$)

   b. Do you think the wedge is operating correctly?

10. Let X be the length in centimeters of a species of fish when caught in the spring. A random sample of 13 observations yielded the sample variance $s^2 = 37.751$. Find a 95% confidence interval for σ. ($\chi^2_{0.025}(12) = 4.404$, $\chi^2_{0.975}(12) = 23.337$).

11. Let X be the length of a male grackle (a type of bird). Suppose $X \sim N(\mu, 4.84)$. Find the sample size that is needed if we are to be 95% confident the maximum error (ie. $z_{\alpha/2}(\sigma/\sqrt{n})$) of the estimate of μ is 0.4. ($z_{0.025} = 1.96$)

12. For a public opinion poll for a close election, let p denote the proportion of votes who favour candidate A. How large a sample should be taken if we want the maximum error of the estimate of p to be equal to

   a. 0.03 with 95% confidence?

   b. 0.02 with 95% confidence?

   c. 0.03 with 90% confidence? ($z_{0.05} = 1.645$).

13. Let $Y_1 < \cdots < Y_5$ be the order statistics of 5 independent observations from an exponential distribution that has a mean of $\theta = 3$.

   a. Find the p.d.f. of the sample minimum Y_1

   b. Compute the probability that $Y_5 < 5$

c. Determine $P(1 < Y_1)$

14. In a clinical trial, let the probability of a successful outcome have a prior distribution that is uniform over $[0, 1]$. Suppose that the first patient has a successful outcome. Find the Bayes estimate of θ that would be obtained for the squared error loss. Also find the Bayes estimate with absolute loss. In both cases, find a 95% posterior probability interval that is symmetric around the Bayes estimate.