620.370 Statistics for Mechanical Engineers — Semester 2, 2009

Homework set 6

Problems to be discussed at next week's tutorial: Quiz 6; 107, 102, 122, 116.

Homework questions

Questions 6.1–6.3 refer to the following information:
A random sample of 100 observations is obtained on the Rayleigh distribution, which has pdf $f(x) = \frac{2x}{\lambda^2} e^{-x^2/\lambda^2}$, $(x > 0)$. The log-likelihood for this sample is given by

$$\ln L = \text{K} - 200 \ln \lambda - \frac{400}{\lambda^2}.$$

6.1 Which statistic produces the 400 in the log-likelihood?
[A.] $2\sum x$; [B.] $2\bar{x}$; [C] $\sum x^2$; [D.] s^2; [E.] $\bar{x}^2$.

6.2 The maximum likelihood estimate of λ is equal to
[A.] $\sqrt{2}$; [B.] 2; [C.] $\sqrt{3}$; [D.] 3; [E.] 4.

6.3 Which one of the following is closest to $\text{se}(\hat{\lambda})$?
[A.] 0.1; [B.] 0.2; [C.] 0.5; [D.] 1; [E.] 10.

Questions 6.4–6.5 refer to the following information:
A random sample of 25 observations on $X \overset{\text{d}}{=} \text{N}(\mu, 1)$ gave $\bar{x} = 2.40$ (and $s = 1.10$).

6.4 A **99**% confidence interval for μ is:
[A.] $(2.01, 2.79)$; [B.] $(1.93, 2.87)$; [C.] $(1.90, 2.90)$; [D.] $(1.88, 2.92)$; [E.] $(1.84, 2.96)$.

6.5 A **99**% prediction interval for X is:
[A.] $(0.48, 4.32)$; [B.] $(0.40, 4.40)$; [C.] $(0.03, 4.77)$; [D.] $(-0.12, 4.92)$; [E.] $(-0.23, 5.03)$.

Data *(also available on the subject web-site in a spreadsheet):*

```
38.0, 32.4, 51.5, 49.3, 47.1, 53.0, 47.7, 52.0, 53.2, 56.7,
67.4, 47.8, 37.6, 37.8, 33.2, 28.4, 36.1, 43.3, 34.8, 24.6,
41.9, 64.5, 31.8, 33.0, 30.6, 34.0, 75.2, 31.5, 32.4, 36.9,
45.8, 32.5, 38.9, 40.6, 45.8, 45.5, 34.1, 53.0, 35.9, 52.8.
```

6.6 (a) Assuming that these data are a random sample from a population X,
i. obtain estimates of the population mean and standard deviation of X;
ii. give an approximate 95% confidence interval for the population mean.

(b) Check whether the distribution of X is Normal using a Normal QQ-plot, i.e. generate a plot of $x_{(k)}$ [vertical axis] vs $\Phi^{-1}(\frac{k}{n+1})$ [horizontal axis], using matlab (or excel, or other). What is your conclusion?
matlab: use
`>> k=1:40; >> fk=k/41; >> ns=norminv(fk); >> xo=sort(x); >> plot(ns,xo)`
excel: order the data in `A1:A50` (from smallest to largest), and put $\frac{1}{41}$, $\frac{2}{41}$, ..., $\frac{40}{41}$ in `B1:B40`; enter `=normsinv(B1)` in `C1`; then plot `A` vs `C` (with `C` on the horizontal-axis).

Quiz 6

Questions Q6.1–Q6.2 refer to the following information:

A Normal QQ-plot for the sample $\{x_1, x_2, \ldots, x_n\}$, plots $x_{(k)}$ vs $\Phi^{-1}(\frac{k}{n+1})$.

Such a QQ-plot for a random sample on $X \stackrel{\mathrm{d}}{=} \mathrm{N}(\mu, \sigma^2)$ passes through the points $(-1, 20)$ and $(2, 50)$.

Q6.1 The estimate of the population mean obtained from the QQ-plot is

[A.] 20; [B.] 25; [C.] 30; [D.] 35; [E.] 40.

Q6.2 The estimate of the population variance obtained from the QQ-plot is

[A.] 10; [B.] 20; [C.] 25; [D.] 50; [E.] 100.

Questions Q6.3–Q6.4 refer to the following information:

A random sample gives the following likelihood function: $L(\theta) = \mathrm{K}\, e^{-5\theta}\theta^{10}$.

Q6.3 The maximum likelihood estimate of θ is

[A.] 1.0; [B.] 1.8; [C.] 2.0; [D.] 3.6; [E.] 7.2.

Q6.4 The standard error of the maximum likelihood estimate of θ is closest to

[A.] 0.4; [B.] 0.6; [C.] 1.2; [D.] 1.6; [E.] 2.5.

Q6.5 In general, the likelihood function, $L(\theta)$, is

[A.] the probability of obtaining the observed sample if the true value of the parameter is θ;

[B.] the likelihood that the true value of the parameter is θ;

[C.] the probability distribution of the parameter;

[D.] the likelihood that the sample is obtained;

[E.] the probability of obtaining the value θ of the parameter given the observed sample.