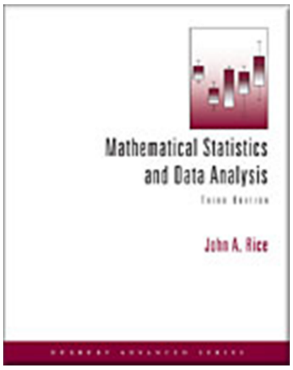
ST122: Probability and Statistics II

Solutions to Selected Problems & Some Extra Materials For:



ISBN 0-534-39942-8

Waleed A. Yousef, Ph.D.,

Human Computer Interaction Lab.,
Computer Science Department,
Faculty of Computers and Information,
Helwan University,
Egypt.

May 6, 2012

Contents

Chapter 6

Distributions Derived from the Normal Distribution

Elaboration on the Independence between $\overline{X}$ and S^2 (advanced topic and requires knowledge of multinormal distribution properties and Linear Algebra)

Suppose that $X_1, \ldots, X_n$ are iid $N\left(\mu, \sigma^2\right)$, then

$$X' = (X_1, \ldots, X_n)' \sim N\left(\mu \mathbf{1}, \sigma^2 \mathbf{I}\right).$$

Consider the transformation

$$Y = AX,$$
$$A = \begin{pmatrix} a\mathbf{1}' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix}.$$

$$\begin{aligned} Y_1 &= a\mathbf{1}'X \\ &= an\overline{X}, \\ Y_i &= b_i^{'}X, \ i = 2, \ldots, n, \\ Y &= N\left(\mu A\mathbf{1}, \sigma^2 AA'\right) \end{aligned}$$

If $\mathbf{1}'b_i = 0$, then $\sigma_{1i} = 0$, $i = 2, \ldots n$, and

$$\Sigma_Y = \begin{pmatrix} \sigma_{11}^2 & \mathbf{0} \\ \mathbf{0} & \Sigma \end{pmatrix},$$

and $Y_1 \left(= \frac{a}{n}\overline{X}\right)$ will be independent of $(Y_2, \ldots, Y_n)'$, which is linear combinations of X. Set

$$b_{ij} = \frac{-1}{n} + \delta_{ij}.$$

That is, b_i is a vector whose all components are $\frac{-1}{n}$ except its ith component. So

$$b_i' X = X_i - \overline{X}, \ i = 2, \ldots, n.$$

Then, $Y_1 \left(= \frac{a}{n}\overline{X}\right)$ is independent of $Y_i \left(= X_i - \overline{X}\right)$, $i = 2, \ldots, n$.Then $\overline{X}$ is independent of $S = S\left(X_2 - \overline{X}, \ldots, X_n - \overline{X}\right)$ (remember that S can be written in terms of $n-1$ differences only since $\Sigma_i \left(X_i - \overline{X}\right) = 0$).
The key reason behind the above result is that the lack of correlation (which can be made up by projection on orthogonal vectors) implies independence in the case of multinormal.

On the other hand, we notice that

$$\begin{aligned} (n-1) S_X &= \sum_i X_i^2 - n\overline{X}^2 \\ &= X'X - n\overline{X}^2. \end{aligned}$$

By orthogonal transformation, $X'X$ will be the same as $Y'Y$, but we can choose Y_1 to be a projection on arbitrary direction. Choose A to be orthogonal with $a = 1/\sqrt{n}$ so that $Y_1^2 = n\overline{X}^2$ and

$$
\begin{aligned}
Y &= N\left(\mu A\mathbf{1}, \sigma^2\mathbf{I}\right), \\
\sum_i X_i^2 &= X'X = Y'Y = \sum_i Y_i^2 \\
(n-1)\,S_X &= X'X - n\overline{X}^2 = Y'Y - Y_1^2 = \sum_{i=2}^{n} Y_i^2 \\
(n-1)\,S_X/\sigma^2 &= \sum_{i=2}^{n} (Y_i/\sigma)^2 \sim \chi^2_{(n-1)}.
\end{aligned}
$$

It is clear that A can be chosen in many other ways to produce different relationships, albiet not meaningful. But there is no something special about $\overline{X}$ and S^2. For example, if we choose A to be orthogonal with

$$
A = \begin{pmatrix} a\,(1,2,\ldots,n)' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix},
$$

then

$$
\begin{aligned}
Y_1 &= a\sum_i iX_i, \\
\sum_i X_i^2 - \left(a\sum_i iX_i\right)^2 &= X'X - \left(a\sum_i iX_i\right)^2 = Y'Y - Y_1^2 = \sum_{i=2}^{n} Y_i^2 \sim \chi^2_{n-1}.
\end{aligned}
$$

Variance of S^2

$$
\begin{aligned}
(n-1)\,S^2/\sigma^2 &\sim \chi^2_{n-1} \\
\operatorname{Var}\left[(n-1)\,S^2/\sigma^2\right] &= 2\,(n-1) && \left(\operatorname{Var}\left[\chi^2_m\right] = 2m\right) \\
\frac{(n-1)^2}{\sigma^4}\operatorname{Var}\left[S^2\right] &= 2\,(n-1) \\
\operatorname{Var}\left[S^2\right] &= \frac{2\sigma^4}{n-1}.
\end{aligned}
$$

For general distribution f_X, after tedious calculations, it can be shown that

$$
\begin{aligned}
\operatorname{Var}\left[S^2\right] &= \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\mu_2^2\right), && (\text{for any } f_X) \\
\mu_i &= \mathrm{E}\left[\left(X_i - \mu_1\right)^2\right]. \\
\operatorname{Var}\left[S^2\right] &= \frac{1}{n}\left(\left(3\sigma^4\right) - \frac{n-3}{n-1}\left(\sigma^2\right)^2\right) && (\text{for } N\left(\mu, \sigma^2\right)) \\
&= 2\frac{\sigma^4}{n-1}
\end{aligned}
$$

Bibliography