ASTP 611-01: Statistical Methods for Astrophysics

Problem Set 5

Assigned 2014 March 7
Problem 1 Due 2014 March 13
Problem 2 Due 2014 March 18

Show your work on all problems! Be sure to give credit to any collaborators, or outside sources used in solving the problems. Note that if using an outside source to do a calculation, you should use it as a reference for the method, and actually carry out the calculation yourself; it's not sufficient to quote the results of a calculation contained in an outside source.

1 Sample Variance

Consider a random sample of size n drawn from a distribution with mean μ and variance σ^2, i.e., random variables $\{X_i|i=1,\dots,n\}$ with $\langle X_i\rangle = \mu$ and $\operatorname{Cov}(X_i,X_j)=\delta_{ij}\sigma^2$

a) What is $\langle X_i - \mu\rangle$?
b) Define $\Sigma^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2$ and calculate $\langle \Sigma^2\rangle$.
c) Define the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and calculate $\langle \overline{X}\rangle$.
d) Calculate $\langle (X_i-\mu)^2\rangle$.
e) Calculate $\left\langle (\overline{X}-\mu)^2\right\rangle$.
f) Calculate $\left\langle (X_i-\mu)(\overline{X}-\mu)\right\rangle$. (Hint: write the sum in the definition of $\overline{X}$ as a sum over j–so as not to repeat the index i–and consider separately the terms in the sum where $j=i$ and $j\neq i$.)
g) Calculate $\left\langle (X_i-\overline{X})^2\right\rangle$. (Hint: write $(X_i-\overline{X})^2 = ([X_i-\mu]-[\overline{X}-\mu])^2$ and use the results of the previous three parts to work out the expectation values of the three terms in the binomial expansion.)
h) Define $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i-\overline{X})^2$ and valculate $\langle S^2\rangle$.

2 Bivariate Normal Distribution

Consider the case of two random variables X_1 and X_2 obeying a bivariate normal distribution $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\sigma}^2_{\mathbf{X}})$ so that $M_{X_1,X_2}(t_1,t_2) = M_{\mathbf{X}}(\mathbf{t}) = \exp(\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\sigma}^2_{\mathbf{X}}\mathbf{t})$ where $\langle X_1\rangle = \mu_1$, $\langle X_2\rangle = \mu_2$, $\operatorname{Var}(X_1)=\sigma_1^2$, $\operatorname{Var}(X_2)=\sigma_2^2$, $\operatorname{Cov}(X_1,X_2)=\rho\sigma_1\sigma_2$, with $-1\le\rho\le 1$, $\sigma_1^2>0$, and $\sigma_2^2>0$.

a) In order to make some of the calculations easier, define $Y_1 = (X_1-\mu_1)/\sigma_1$ and $Y_2 = (X_2-\mu_2)/\sigma_2$. Show that $\langle Y_1\rangle = 0 = \langle Y_2\rangle$, $\operatorname{Var}(Y_1) = 1 = \operatorname{Var}(Y_2)$, and $\operatorname{Cov}(Y_1,Y_2)=\rho$, and that $\mathbf{Y}\sim N_2(\mathbf{0},\boldsymbol{\sigma}^2_{\mathbf{Y}})$, where

$$\boldsymbol{\sigma}^2_{\mathbf{Y}} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{2.1}$$

b) Show that

$$\mathbf{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \tag{2.2}$$

are orthonormal eigenvectors of $\boldsymbol{\sigma}^2_{\mathbf{Y}}$, and find the corresponding eigenvalues λ_1 and λ_2. Verify that $\boldsymbol{\sigma}^2_{\mathbf{Y}} = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^{\mathrm{T}} + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^{\mathrm{T}}$.

c) Take the determinant $\det \boldsymbol{\sigma}^2_{\mathbf{Y}}$ and verify that $\det \boldsymbol{\sigma}^2_{\mathbf{Y}} = \lambda_1 \lambda_2$, and that this determinant is positive unless $\rho = 1$ or $\rho = -1$

d) Assuming $-1 < \rho < 1$, find the matrix inverse $\boldsymbol{\sigma}^{-2}_{\mathbf{Y}} = (\boldsymbol{\sigma}^2_{\mathbf{Y}})^{-1}$ and verify that $\boldsymbol{\sigma}^{-2}_{\mathbf{Y}} = (\lambda_1)^{-1} \mathbf{v}_1 \mathbf{v}_1^{\mathrm{T}} + (\lambda_2)^{-1} \mathbf{v}_2 \mathbf{v}_2^{\mathrm{T}}$.

e) If $-1 < \rho < 1$, the pdf for $\mathbf{Y}$ is

$$f(\mathbf{y}) = \frac{1}{\sqrt{\det(2\pi \boldsymbol{\sigma}^2_{\mathbf{Y}})}} \exp\left(-\frac{1}{2} (\mathbf{y} - \mathbf{0})^{\mathrm{T}} \boldsymbol{\sigma}^{-2}_{\mathbf{Y}} (\mathbf{y} - \mathbf{0}) \right) \; ; \tag{2.3}$$

write this explicitly as a joint pdf $f(y_1, y_2)$ without using any matrix expressions.

f) Define $Y_+ = (Y_1 + Y_2)/\sqrt{2}$ and $Y_- = (Y_1 - Y_2)/\sqrt{2}$, and perform a transformation of variables on your $f(y_1, y_2)$ to find the joint pdf $f(y_+, y_-)$. Use this pdf to verify that Y_+ and Y_- are independent Gaussian random variables. What are the parameters of the Gaussian distributions for Y_+ and Y_-?

g) If we marginalize over Y_2, we can get $f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2$. Evaluate this integral and verify by inspection of the pdf that Y_1 is a standard normal random variable.

h) If, on the other hand, we assume a value y_2 for Y_2, we can define the conditional pdf

$$f_{1|2}(y_1|y_2) = f(y_1|Y_2 = y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} \tag{2.4}$$

Using the form of $f(y_1, y_2)$ from part e), and the marginal pdf $f_2(y_2) = \frac{1}{\sqrt{2\pi}} e^{-(y_2)^2/2}$ [by analogy to the result of part f)], work out this conditional pdf, and verify that it is a Gaussian.

i) A common pitfall when dealing with correlated errors is to choose the most likely value for one parameter and then consider the width of the distribution for another parameter, assuming that most likely value, which leads to an underestimate of the errors. In this case, that would mean assuming $y_2 = 0$ and using the conditional pdf $f_{1|2}(y_1|0) = f(y_1|Y_2 = 0)$ rather than the marginalized pdf $f_1(y_1)$. Compare the width of these two Gaussians.

j) The combination $\chi^2(\mathbf{Y}) = (\mathbf{Y} - \mathbf{0})^{\mathrm{T}} \boldsymbol{\sigma}^{-2}_{\mathbf{Y}} (\mathbf{Y} - \mathbf{0})$ obeys a chi-squared distribution with two degrees of freedom. Write $\chi^2(y_1, y_2)$ as an explicit function of y_1 and y_2 without any matrix expressions. What shape does a curve of constant $\chi^2(y_1, y_2)$ trace in the (y_1, y_2) plane?

k) For $\rho = \frac{1}{2}$, use the plotting program of your choice to make a plot in the (y_1, y_2) plane with all of the following shown on it:

   i) The curve $\chi^2(y_1, y_2) = 1$
   ii) Error bars centered on $y_1 = 0$ extending to plus and minus one standard deviation of the marginalized pdf $f_1(y_1)$
   iii) Error bars centered on $y_1 = 0$ extending to plus and minus one standard deviation of the conditional pdf $f_{1|2}(y_1|0) = f(y_1|Y_2 = 0)$