Statistical methods – an introduction (SS 2016)

Problem set 4

Problem 1 [7 points] *Transformation of variables*

a) Calculate the transformation law $y(x)$ to generate random variables distributed according to a pdf $g(y) = y/50$ in the range $y \in [0, 10]$ (and $g(y) = 0$ else), by means of a random number generator. A typical RNG creates uniformly distributed variates in [0,1], according to a pdf

$$f(x) = 1 \text{ if } x \in [0, 1] \qquad \text{and} \qquad f(x) = 0 \text{ else.}$$

b) The variates y are transformed via $u(y) = (y-5)^2$. What is the resulting distribution $h(u)$? If you are not able to solve this problem, here is the result:

$$h(u) = \frac{1}{10\sqrt{u}} \text{ if } u \in [0, 25] \qquad \text{and} \qquad h(u) = 0 \text{ else.}$$

Convince yourself that this distribution is normalized.

c) Calculate, from first principles, the transformation law $u'(y)$ which is required to generate random variables distributed according to $h(u)$, from variates distributed according to $g(y)$. Compare the transformation laws $u(y)$ and $u'(y)$, and provide some conclusions.

d) Check your results, by comparing the distributions of numerically generated random numbers according to $h(u)$, after transforming variates following $g(y)$ via $u(y)$ and $u'(y)$. Compare also with the analytical distribution. Proceed in analogy to page 66 of the script.

IDL: Make use of routine `my_histogram.pro` from *problem set 2*. For the RNG, use `randomu`. Use the keywords `\plot` and `\norm` when calling `my_histogram`. To allow for a clearer plot, use an ordinate-range of [0,0.3]: `!y.range=[0.,0.3]`

PYTHON: Make use of routine `my_hist.py` (from the homepage). For the RNG, use `np.random.uniform`. Use the keywords `norm` when calling `my_hist.hist`, and `oplot` and `color` when overplotting the 2nd histogram. To allow for a clearer plot, use an ordinate-range of [0,0.3]: `plt.ylim(0.,0.3)`.

If you are not able to create variates following $g(y)$, i.e., you could not solve problem a), please contact your supervisor who will tell you the transformation law.

Problem 2 [5 points] *Correlation*

a) Prove that the correlation coefficient $\rho(x, y) = \pm 1$ if $y = a + bx$ (see page 78 of manuscript).

b) Now let $y = a + bx + cx^2$. Calculate $\rho(x, y)$, if the distribution of x is symmetric about '0'.

c) Calculate $\rho(x, y)$ for the above example, when x is uniformly distributed in [-1,1], and $a = 2, b = 3, c = 4$. Test your result via the function `correlate` (IDL) or `np.corrcoef` (PYTHON), based on a sufficiently large sample of uniformly distributed random numbers. What do you expect (no calculation required, just argue!) for the case when x is uniformly distributed within [0,1]? Test your expectation by again using `correlate/np.corrcoef`.

Have fun, and much success!