DSE 210: Probability and Statistics using Python

Worksheet 6 — Modeling dependence between variables

1. Would you expect the following pairs of random variables to be uncorrelated, positively correlated, or negatively correlated?
   (a) The amount of rainfall on a given day and the amount of rainfall the following day.
   (b) The number of people at the beach on a given day and the number of people skiing that day.
   (c) A person's age and social security number.

2. *Correlation and independence.* Random variables X and Y take values in $\{-1, 0, 1\}$. Their joint distribution is given in the following table:

		Y		
		-1	0	1
	-1	1/8	0	1/8
X	0	0	1/2	0
	1	1/8	0	1/8

   (a) Are X and Y independent?
   (b) Compute the correlation coefficient between X and Y.

3. *Covariance and correlation.* Random variable X has mean zero and standard deviation 10. Random variable Y is defined by $Y = 2X$.
   (a) What is the covariance between X and Y?
   (b) What is the correlation coefficient between X and Y?

4. Each of the following scenarios describes a joint distribution (x, y). In each case, give the parameters of the (unique) bivariate Gaussian that satisfies these properties.
   (a) x has mean 2 and standard deviation 1, y has mean 4 and standard deviation 0.5, and the correlation between x and y is -0.5.
   (b) x has mean 1 and standard deviation 1, and y is equal to x.

5. *More bivariate Gaussians.* Roughly sketch the shapes of the following Gaussians $N(\mu, \Sigma)$. For each, you only need to show a representative contour line which is qualitatively accurate (has approximately the right orientation, for instance).
   (a) $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$
   (b) $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$

6. For each of the two Gaussians in the previous problem, check your answer using Python: draw 100 random samples from that Gaussian and plot them.

7. *Qualitative appraisal of Gaussian parameters.* A bivariate Gaussian has covariance matrix $\begin{pmatrix} p & q \\ q & r \end{pmatrix}$. Give precise characterizations, in terms of p, q, r, of when the following are true.

   (a) The two variables are negatively correlated.

   (b) The two variables are uncorrelated.

   (c) One variable is a linear function of the other.

   (d) The second variable is a constant (i.e. always takes the same value).