Tutorial 5 Problems

for discussion on Week 48

Main discrete distributions

1. Let ξ be a discrete random variable uniformly distributed over numbers $1, 2, \dots, n$. Find its expectation and the variance.
2. A r.v. ξ is the number of trials to get the first 'success' in the series of independent Bernoulli trials with success probability p. Find its distribution which is called a *Geometric districution* $\mathrm{Geom}(p)$ with parameter $p \in (0, 1)$, and find its expectation and variance.
3. Problem 3.8.6 from GS[1]
4. Problem 3.11.14 from GS
5. Problem 3.11.17 from GS

Main continuous distributions

1. Let $\eta \sim \mathrm{Unif}(a, b)$. Find the distribution of $c\eta + d$ for some $c, d \in \mathbb{R}$.
2. Compute the expected value and the variance of the Exponential $\mathrm{Exp}(\lambda)$ distribution which has the density $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ and 0 otherwise.
3. Show that the variance of the Normal $\mathcal{N}(m, \sigma^2)$ distribution is σ^2.
4. Finish the proof started at the lecture that for any $a \in \mathbb{R}, b \in (0, \infty)$ and a r.v. $\xi \sim \mathcal{N}(m, \sigma^2)$, the random variable $a\xi + b$ is Normal. What are its parameters?
5. Let ξ is absolutely continuous with pdf $f_\xi(x)$. Show that $\zeta = a\xi + b$ is also absolutely continuous and express its pdf in terms of $f_\xi(x)$.

[1] As usual, GS stands for the course book: Geoffrey Grimmett and David Stirzaker, *Probability and Random Processes*, Oxford University Press, 3rd edition, 2001. Problem x.y.z means Problem z for Section x.y in this book

6. Problem 4.7.8 from GS
7. A random point (ξ, η) is uniformly distributed in the square $\{(x, y) : |x| + |y| \le \sqrt{2}\}$. Find
   (a) the joint c.d.f. of the pair (ξ, η);
   (b) marginal distribution of ξ;
   (c) $\mathbf{E}\,\xi$ and $\mathbf{var}\,\xi$;
   (d) covariance $\mathbf{cov}(\xi, \eta)$ and the correlation coefficient $\mathbf{cor}(\xi, \eta)$.
   (e) Check if ξ and η are independent.
8. Problem 4.7.1 from GS
9. The *covariance* of two random variables is defined as

$$\mathbf{cov}(\xi_1, \xi_2) = \mathbf{E}[(\xi_1 - \mathbf{E}\,\xi_1)(\xi_2 - \mathbf{E}\,\xi_2)]\,.$$

   Show that

$$\mathbf{cov}(\xi_1, \xi_2) = \mathbf{E}[\xi_1\xi_2] - \mathbf{E}\,\xi_1\,\mathbf{E}\,\xi_2$$

   and establish the following properties of the covariance: whatever the constants $a, c \in \mathbb{R}$,

   1. $\mathbf{cov}(\xi_1, \xi_2) = \mathbf{cov}(\xi_2, \xi_1)$;
   2. $\mathbf{cov}(\xi_1, c) = 0$;
   3. $\mathbf{cov}(a\xi_1, \xi_2) = a\,\mathbf{cov}(\xi_1, \xi_2)$;
   4. $\mathbf{cov}(\xi_1 + \xi_2, \xi_3) = \mathbf{cov}(\xi_1, \xi_3) + \mathbf{cov}(\xi_2, \xi_3)$;
   5. $\mathbf{cov}(\xi_1, \xi_2) = 0, \quad$ if ξ_1, ξ_2 are independent.

10. The (Pearson) *correlation coefficient* of two random variables is defined as

$$\mathbf{cor}(\xi_1, \xi_2) = \frac{\mathbf{cov}(\xi_1, \xi_2)}{\mathbf{var}\,\xi_1\;\mathbf{var}\,\xi_2}\,.$$

   Based on the previous exercise, show the following properties of the correlation coefficient: whatever the constants $a, c \in \mathbb{R}$,

1. $\mathbf{cor}(\xi_1, \xi_2) = \mathbf{cor}(\xi_2, \xi_1)$;
2. $\mathbf{cor}(\xi_1, c) = 0$;
3. $\mathbf{cor}(a\xi_1, \xi_2) = \operatorname{sgn}(a)\ \mathbf{cor}(\xi_1, \xi_2)$, where
$$\operatorname{sgn}(a) = \begin{cases} 1, & \text{if } a > 0 \\ 0, & \text{if } a = 0 \\ -1, & \text{if } a < 0; \end{cases}$$
4. $\mathbf{cor}(\xi_1, \xi_2) = 0$, if ξ_1, ξ_2 are independent.

NB. The last two exercises show that the covariance have the same properties as the scalar product in $\mathbb{R}^d$ defined for two vectors $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ by means of $(x, y) = \sum_{i=1}^d x_i y_i$. Notice that $(x, x) = \|x\|^2$ is the squared length (norm) of x. Similarly, covariance can be considered as a scalar product in the space $\mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$ of centred (i.e. having expectation 0) random variables ξ such that their norm $\|\xi\| \stackrel{\text{def}}{=} \sqrt{\mathbf{cov}(\xi, \xi)} = \sqrt{\mathbf{E}\,\xi^2}$ is finite. Since $(x, y)/(\|x\|\|y\|)$ is the cosine of the angle between x and y in $\mathbb{R}^d$, the correlation coefficient has also a meaning of the cosine of the angle between ξ and η. This implies that it is always between -1 and 1! When it is 1, the angle is 0 so that ξ and η are co-linear pointing in the same direction, so that $\eta = a\xi$ for some $a > 0$. Similarly, when the correlation is -1, they point in the opposite directions (the angle is π) so that $\eta = a\xi$ with a negative a. Another interesting observation is that uncorrelated centred random variables are orthogonal vectors in $\mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$.

11. Problem 3.11.16 from GS (*uncorrelated* means the correlation coefficient, and so is the covariance, is 0)