Solutions to Exercise sheet 2

Exercise 1: We roll three dice. The random variables X_1, X_2, X_3 with values in $\{1, 2, 3, 4, 5, 6\}$ encode the results of the three rolls. What is the cdf of the maximum of the three rolls $X = \max\{X_1, X_2, X_3\}$?

Solution for Exercise 1: We compute the cdf F_X of X. By definition

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \le x) \\ &\overset{(1)}{=} \mathbb{P}(X_1 \le x \text{ and } X_2 \le x \text{ and } X_3 \le x) \\ &\overset{(2)}{=} \mathbb{P}(X_1 \le \lfloor x \rfloor)\mathbb{P}(X_2 \le \lfloor x \rfloor)\mathbb{P}(X_3 \le \lfloor x \rfloor) \\ &\overset{(3)}{=} \frac{\lfloor x \rfloor^3}{6^3}, \end{aligned}$$

for all $x \in [1, 6]$, where $\lfloor x \rfloor$ rounds down any real number to the next lower integer (for integers $x = \lfloor x \rfloor$). Here we used in (1) that $\{\max\{X_1, X_2, X_3\} \le x\} = \{X_1 \le x, X_2 \le x, X_3 \le x\}$. In (2) we used the product rule for probabilities when random experiments are independent (Of course the rolls of the three dice here are independent of each other.) and that the dice rolls X_i can only have integer values. In (3) we used that

$$\mathbb{P}(X_j \le i) = \frac{i}{6}$$

for all $i = 1, 2, 3, 4, 5, 6$ and $j = 1, 2, 3$ since a dice roll is uniformly distributed on $\{1, 2, 3, 4, 5, 6\}$. For values $x < 0$ we have $F_X(x) = 0$ because X cannot take any of these values with non-zero probability. For $x > 6$ we have $F_X(x) = 1$ because in this case $\{X \le x\}$ happens with probability 1.

Exercise 2: Let $X \sim N(0, v)$ with $v > 0$. What is the pdf of X^2?

Solution for Exercise 2: Let $A \subset [0, \infty)$ be an interval. By definition of the pdf we want to know whether

$$\mathbb{P}(X^2 \in A) = \int_A \rho(x)\mathrm{d}x$$

for some density function $\rho = \rho_{X^2}$. We determine this function by computing

$$\mathbb{P}(X^2 \in A) \overset{(1)}{=} \mathbb{P}(X \in \sqrt{A}) \overset{(2)}{=} \frac{1}{\sqrt{2\pi v}} \int_{\sqrt{A}} \mathrm{e}^{-\frac{x^2}{2v}} \mathrm{d}x \overset{(3)}{=} \frac{1}{\sqrt{2\pi v}} \int_A \mathrm{e}^{-\frac{y}{2v}} \frac{1}{2\sqrt{y}} \mathrm{d}y\,.$$

In (1) we used that

$$\{X^2 \in A\} = \{\omega \in \Omega : X(\omega)^2 \in A\} = \{\omega \in \Omega : X(\omega) \in \sqrt{A}\} = \{X \in \sqrt{A}\}\,,$$

where $\sqrt{A} = \{\sqrt{x} : x \in A\} \subset [0, \infty)$ and Ω is the underlying probability space. We compare to the definition of the pdf and see that the pdf of X^2 is

$$\rho_{X^2}(y) = \frac{\mathrm{e}^{-\frac{y}{2v}}}{2\sqrt{2\pi v y}}\,.$$

Exercise 3: Let X be a continuous random vector with pdf $\rho : \mathbb{R}^d \to \mathbb{R}$ and $\alpha \in \mathbb{R}$, $\mu \in \mathbb{R}^d$ with $\alpha \neq 0$. What is the pdf of $\alpha X + \mu$?

Solution for Exercise 3: We imitate the proof of the transformation theorem for probability density functions (or apply the theorem). As in Exercise 2 we will use the definition of the pdf and therefore we compute for some $A = A_1 \times \cdots \times A_d$ the probability

$$\begin{aligned}
\mathbb{P}(\alpha X + \mu \in A) &= \mathbb{P}\Big(X \in \frac{A-\mu}{\alpha}\Big) \\
&\stackrel{(1)}{=} \int_{\mathbb{R}^d} \rho_X(x) \mathbb{1}\Big(x \in \frac{A-\mu}{\alpha}\Big)\mathrm{d}x \\
&\stackrel{(2)}{=} \int_{\mathbb{R}^d} \rho_X(x)\mathbb{1}(\alpha x + \mu \in A)\mathrm{d}x \\
&\stackrel{(3)}{=} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \rho_X(x_1, \dots, x_d)\mathbb{1}((\alpha x_1 + \mu_1, \dots, \alpha x_1 + \mu_d) \in A)\mathrm{d}x_1 \dots \mathrm{d}x_d \\
&\stackrel{(4)}{=} \frac{1}{|\alpha|^d} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \rho_X\Big(\frac{y_1 - \mu_1}{\alpha}, \dots, \frac{y_d - \mu_d}{\alpha}\Big)\mathbb{1}((y_1, \dots, y_d) \in A)\mathrm{d}y_1 \dots \mathrm{d}y_d \\
&\stackrel{(5)}{=} |\alpha|^{-d} \int \rho_X\Big(\frac{y-\mu}{\alpha}\Big)\mathbb{1}(y \in A)\mathrm{d}y\,.
\end{aligned}$$

In (1) we used the definition of the pdf for X, namely that $\mathbb{P}(X \in B) = \int_B \rho_X(x)\mathrm{d}x = \int_{\mathbb{R}^d} \rho_X(x)\mathbb{1}_B(x)\mathrm{d}x$ and the intuitive notation $\mathbb{1}_B(x) = \mathbb{1}(x \in B)$. In (2) we rewrite the indicator function. In (3) we write out the d-dimensional integral as iterated 1-dimensional integrals (as defined in the lecture). In (4) we change variables to $y_i = \alpha x_i + \mu_i$ in each 1-dimensional integral and use the substitution rule. In (5) we rewrite the integral back into a d-dimensional integral. Comparing to the definition of the pdf we conclude that $\rho_{\alpha X + \mu}(y) = |\alpha|^{-d}\rho_X\big(\frac{y-\mu}{\alpha}\big)$.

Here are some more details of performing the substitution in step (4). Note how the substitution rule in dimension 1 behaves when we shift and rescale the variable of a function f (that we can integrate). Here is the behaviour for $\beta > 0$ and $\gamma \in \mathbb{R}$:

$$\int_{\mathbb{R}} f(\beta t + \gamma)\mathrm{d}t = \int_{-\infty}^{\infty} f(\beta t + \gamma)\mathrm{d}t == \frac{1}{\beta}\int_{-\infty}^{\infty} f(s)\mathrm{d}s = \frac{1}{\beta}\int_{\mathbb{R}} f(t)\mathrm{d}t\,.$$

And here it is for $\beta < 0$ and $\gamma \in \mathbb{R}$:

$$\int_{\mathbb{R}} f(\beta t + \gamma)\mathrm{d}t = \int_{-\infty}^{\infty} f(\beta t + \gamma)\mathrm{d}t = \frac{1}{\beta}\int_{\infty}^{-\infty} f(s)\mathrm{d}s = -\frac{1}{\beta}\int_{-\infty}^{\infty} f(t)\mathrm{d}t = \frac{1}{|\beta|}\int_{\mathbb{R}} f(t)\mathrm{d}t\,.$$

Thus, in general we have

$$\int_{\mathbb{R}} f(\beta t + \gamma)\mathrm{d}t = \frac{1}{|\beta|}\int_{\mathbb{R}} f(t)\mathrm{d}t$$

for any $\beta \neq 0$ and $\gamma \in \mathbb{R}$. For $\beta = 1$ we see that the integral is invariant under shifts.

The factor $|\alpha|^{-d}$ in (4) this stems from performing d one dimensional rescalings while the shift by μ does gives rise to a trivial Jacobian (i.e. no extra factor or function in the integrand).

Exercise 4: Let X be a standard normal random variable. Use the Markov inequality to show the following estimate on large deviations

$$\mathrm{e}^{\alpha\xi^2}\mathbb{P}(|X| \geq \xi) \to 0\,, \qquad \xi \to \infty\,.$$

for all $\alpha < 1/2$.

Solution for Exercise 4: We apply the Markov inequality to the function $f(\xi) = \mathrm{e}^{\alpha\xi^2}$ and find

$$\mathbb{P}(|X| \geq \xi) \leq \frac{\mathbb{E}\mathrm{e}^{\alpha X^2}}{\mathrm{e}^{\alpha\xi^2}} = \frac{\mathrm{e}^{-\alpha\xi^2}}{\sqrt{2\pi}}\int \mathrm{e}^{-(1/2-\alpha)x^2}\mathrm{d}x\,. \tag{0.1}$$

For the equality we used the definition of a standard normal random variable and how to compute the expectation of a continuous randpom variable, i.e. we used

$$\mathbb{E} f(X) = \int_{\mathbb{R}} f(x)\rho_X(x)\mathrm{d}x\,, \qquad \rho_X(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,.$$

The integral on the right hand side of (0.1) is still finite if $\alpha < 1/2$ as we have seen in the lecture when we integrated Gaussian densities.

Exercise 5: Compute expectation and variance of a real random variable X with the following property (each item is a different exercise):

1. X has cdf $F_X(x) = 1 - \mathrm{e}^{-\alpha x}$ for $x \geq 0$ with some constant $\alpha > 0$.
2. $X = Y^k$, where Y is uniformly distributed on $[0,1]$ and $k \in \mathbb{N}$.
3. $X \sim \mathrm{Poi}(\lambda)$ for some $\lambda > 0$.

Solution for Exercise 5: Solution of 1: Since the cdf of X does not have jumps and is zero on negative numbers (since $F_X(0) = 0$ and F_X is monotonously increasing with non-negative values) we suspect that X must be a continuous random variable. To compute its pdf we differentiate F_X (compare Homework 1 on Assignment Sheet 2). Thus,

$$\rho_X(x) = F_X'(x) = \alpha\,\mathrm{e}^{-\alpha x},$$

i.e. $X \sim \mathrm{Exp}(\alpha)$ (as defined in the lecture). Now we compute the first two moments. The expectation is

$$\mathbb{E}\,X \overset{(1)}{=} \int_0^\infty x\,\alpha\,\mathrm{e}^{-\alpha x}\mathrm{d}x \overset{(2)}{=} -\alpha\frac{\mathrm{d}}{\mathrm{d}\alpha}\int_0^\infty \mathrm{e}^{-\alpha x}\mathrm{d}x \overset{(3)}{=} -\alpha\frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{1}{\alpha} \overset{(4)}{=} \frac{1}{\alpha}\,.$$

In (1) we used the formula for computing expectations of continuous random variables. In (2) we use a useful trick that gets rid of the factor x in the integrand, namely to write $x\mathrm{e}^{-\alpha x} = -\frac{\mathrm{d}}{\mathrm{d}\alpha}\mathrm{e}^{-\alpha x}$ and then to pull the derivative out of the integral. Instead of that we could have simply used integration by parts, but this trick is still useful to keep in mind. In (3) we integrated the exponential

$$\int_0^\infty \mathrm{e}^{-\alpha x}\mathrm{d}x = -\frac{1}{\alpha}\mathrm{e}^{-\alpha x}|_0^\infty = -\frac{1}{\alpha}(0-1) = \frac{1}{\alpha}\,.$$

In (5) we differentiate $\frac{\mathrm{d}}{\mathrm{d}\alpha}\alpha^{-1} = -\alpha^{-2}$.

Now that we have computed the expectation we will determine the variance. For that we compute the second moment

$$\mathbb{E}\,X^2 = \int_0^\infty x^2\,\alpha\,\mathrm{e}^{-\alpha x}\mathrm{d}x = \alpha\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\int_0^\infty \mathrm{e}^{-\alpha x}\mathrm{d}x = \alpha\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\frac{1}{\alpha} = \frac{2}{\alpha^2}\,.$$

Note that we used the same trick with differentiating the exponent and in this way the calculation did not become any longer (except that we have to take two derivatives at the end now). We could have used integration by part twice instead, but that is a slightly longer and messier calculation. Now we compute the variance

$$\mathrm{Var(X)} = \mathbb{E}\mathrm{X}^2 - (\mathbb{E}\mathrm{X})^2 = \frac{2}{\alpha^2} - \frac{1}{\alpha^2} = \frac{1}{\alpha^2}\,.$$

Solution of 2: The pdf of a uniformly distributed random variable on $[0,1]$ is $\rho_Y(y) = \mathbb{1}(y \in [0,1])$.

Now we compute the expectation (using again the definition of expectation for continuous random variables)

$$\mathbb{E}\,X = \mathbb{E}Y^k = \int_0^1 y^k\mathrm{d}y = \frac{1}{k+1}\,.$$

We recycle this calculation to compute the second moment by replacing k with $2k$ and find

$$\mathbb{E}X^2 = \mathbb{E}Y^{2k} = \frac{1}{2k+1}\,.$$

For the variance we find

$$\mathrm{Var}(X) = \mathbb{E}\,X^2 - (\mathbb{E}\,X)^2 = \frac{1}{2k+1} - \frac{1}{(k+1)^2}.$$

Solution of 3: Here the random variable is discrete with values in non-negative integers. The pmf of X is $p_X(x) = \frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}$. Using this we compute the expectation via its definition for discrete random variables

$$\mathbb{E}\,X = \sum_{x=0}^{\infty} x\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda} \overset{(1)}{=} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!}\mathrm{e}^{-\lambda} \overset{(2)}{=} \lambda\mathrm{e}^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} \overset{(3)}{=} \lambda\,.$$

In (1) we cancelled an x and removed $x = 0$ from the sum (since the corresponding summand is zero). In (2) we rename the index in the sum, shifting it by 1. In (3) we use the exponential series $\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = \mathrm{e}^{\lambda}$. Now we turn to the second moment

$$\begin{aligned}\mathbb{E}\,X^2 &= \sum_{x=0}^{\infty} x^2\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda} \overset{(1)}{=} \sum_{x=1}^{\infty} x\frac{\lambda^x}{(x-1)!}\mathrm{e}^{-\lambda} \overset{(2)}{=} \mathrm{e}^{-\lambda}\left(\sum_{x=1}^{\infty}(x-1)\frac{\lambda^x}{(x-1)!} + \sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!}\right)\\ &\overset{(3)}{=} \mathrm{e}^{-\lambda}\left(\lambda^2\mathrm{e}^{\lambda} + \lambda\mathrm{e}^{\lambda}\right) = \lambda^2 + \lambda\,.\end{aligned}$$

In (1) we cancelled an x. Since we would like to cancel $x-1$ as well we write $x = x-1+1$ and split the series into two in (5). In (3) we shift the summation index and proceed as we did for the expectation.

Finally, putting expectation and second moment together we arrive at the variance

$$\mathrm{Var(X)} = \mathbb{E}\mathrm{X}^2 - (\mathbb{E}\mathrm{X})^2 = \lambda\,.$$

Exercise 6: Use the fact that the Gaussian density

$$\rho(x) = \frac{1}{\sqrt{2\pi v}}\mathrm{e}^{-\frac{(x-a)^2}{2v}}$$

is a probability density, i.e. that $\int\rho(x)\mathrm{d}x = 1$, for any $a\in\mathbb{R}$ and $v > 0$ to conclude

$$\int \mathrm{e}^{-\frac{\alpha}{2}x^2+\beta x}\mathrm{d}x = \sqrt{\frac{2\pi}{\alpha}}\,\mathrm{e}^{\frac{\beta^2}{2\alpha}}$$

for any $\beta\in\mathbb{R}$ and $\alpha > 0$. Furthermore, compute expectation and variance of a random variable X with distribution $N(a, v)$.

Solution for Exercise 6: We compute then integral

$$\int \mathrm{e}^{-\frac{\alpha}{2}x^2+\beta x}\mathrm{d}x \overset{(1)}{=} \frac{1}{\sqrt{\alpha}}\int \mathrm{e}^{-\frac{x^2}{2}+\beta x/\sqrt{\alpha}}\mathrm{d}x \overset{(2)}{=} \frac{1}{\sqrt{\alpha}}\int \mathrm{e}^{-\frac{1}{2}(x-\beta/\sqrt{\alpha})^2+\frac{\beta^2}{2\alpha}}\mathrm{d}x \overset{(3)}{=} \frac{\sqrt{2\pi}\mathrm{e}^{\frac{\beta^2}{2\alpha}}}{\sqrt{\alpha}}\int\frac{\mathrm{e}^{-\frac{1}{2}(x-\beta/\sqrt{\alpha})^2}}{\sqrt{2\pi}}\mathrm{d}x\,. \tag{0.2}$$

In (1) we used the substitution rule for the change of variables $x \to \frac{x}{\sqrt{\alpha}}$ (or put differently we choose a new variable $y = \sqrt{\alpha}x$ and afterwards rename the integration variables back to x). In (2) we complete the square in the exponent. In (3) we pulled out a factor after factorising with the formula $\mathrm{e}^{x+y} = \mathrm{e}^x\mathrm{e}^y$.

The integral remaining on the right hand side of (0.2) can be simplified by shifting, namely

$$\int\frac{\mathrm{e}^{-\frac{1}{2}(x-\beta/\sqrt{\alpha})^2}}{\sqrt{2\pi}}\mathrm{d}x = \int\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\mathrm{d}x\,.$$

But this was computed in the lecture. It is the integral over the pdf of a standard normal random variable and therefore equals 1.

Now we compute the expectation of X via

$$\begin{aligned}\mathbb{E}X &\overset{(1)}{=} \frac{1}{\sqrt{2\pi v}}\int x\,\mathrm{e}^{-\frac{(x-a)^2}{2v}}\,\mathrm{d}x\\ &\overset{(2)}{=} \frac{1}{\sqrt{2\pi v}}\int (x+a)\,\mathrm{e}^{-\frac{x^2}{2v}}\,\mathrm{d}x\\ &\overset{(3)}{=} \frac{1}{\sqrt{2\pi v}}\int x\,\mathrm{e}^{-\frac{x^2}{2v}}\,\mathrm{d}x + \frac{a}{\sqrt{2\pi v}}\int \mathrm{e}^{-\frac{x^2}{2v}}\,\mathrm{d}x\\ &\overset{(4)}{=} a\,.\end{aligned}$$

In (1) we used the definition of the expectation for continuous random variables. In (2) we used a shift $x \to x + a$ of the integration variable. In (3) we split the integral into two summand. In (4) we use that

$$\int x\,\mathrm{e}^{-\frac{x^2}{2v}}\,\mathrm{d}x = 0\,, \qquad \frac{1}{\sqrt{2\pi v}}\int \mathrm{e}^{-\frac{x^2}{2v}}\,\mathrm{d}x = 1\,.$$

The first identity holds because the integrand is antisymmetric and thus integrates to zero. The second identity is the normalisation of the Gaussian pdf. We proceed similarly for the variance

$$\begin{aligned}\mathrm{Var}\,\mathrm{X} &= \frac{1}{\sqrt{2\pi v}}\int (x-a)^2\,\mathrm{e}^{-\frac{(x-a)^2}{2v}}\,\mathrm{d}x\\ &= \frac{1}{\sqrt{2\pi v}}\int x^2\,\mathrm{e}^{-\frac{x^2}{2v}}\,\mathrm{d}x\\ &= \frac{v}{\sqrt{2\pi}}\int x^2\,\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x\\ &= v\,.\end{aligned}$$

To compute the integral in the last step we can use e.g. our trick of differentiating with respect to the exponent again

$$\frac{1}{\sqrt{2\pi}}\int x^2\,\mathrm{e}^{-\frac{\alpha}{2}x^2}\,\mathrm{d}x = -\frac{2}{\sqrt{2\pi}}\frac{\mathrm{d}}{\mathrm{d}\alpha}\int \mathrm{e}^{-\frac{\alpha}{2}x^2}\,\mathrm{d}x = -\frac{2}{\sqrt{2\pi}}\frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{\sqrt{2\pi}}{\sqrt{\alpha}} = \alpha^{-3/2}\,.$$

For $\alpha = 1$ this evaluates to $\alpha^{-3/2} = 1$.

We could have also used the substitution rule with $x \to \sqrt{x}$ and then the facts about the Γ-function from the lecture (that you prove on the homework sheet).