MSIAM M1: Probability and Statistics

Solutions to [the most difficult] exercises

1 TD 1

Exercise 1.2. *Provide an example of asymmetric density with $\alpha = 0$.*

Proof. Consider a discrete random variable X, such that $\mathbb{P}(X = 2) = \mathbb{P}(X = -1) = \frac{1}{4}, \mathbb{P}(X = \sqrt{7}) = \frac{1}{4\sqrt{7}}$, and $\mathbb{P}(X = 0) = \frac{1}{2} - \frac{1}{4\sqrt{7}}$. First, note that this r.v. is asymmetric. Indeed, $\mathbb{E}[X] = \frac{1}{2} - \frac{1}{4} - \sqrt{7}\frac{1}{4\sqrt{7}} = 0$. However,

$$F_X(0) = \frac{3}{4} \quad \neq \quad 1 - F_X(0) = \frac{1}{4}.$$

Now, compute the third moment of X:

$$\mathbb{E}[X^3] = 2 - \frac{1}{4} - 7\sqrt{7}\frac{1}{4\sqrt{7}} = 0.$$

Thus, the skewness parameter is 0. □

Exercise 1.4. *Let (ξ_n) and (η_n) be two sequences of r.v. Prove the following statements:*

1°. *If $a \in \mathbb{R}$ is a constant, then when $n \to \infty$:*

$$\xi_n \xrightarrow{D} a \quad \Leftrightarrow \quad \xi_n \xrightarrow{P} a$$

2°. *(**Slutsky's theorem**.) If $\xi_n \xrightarrow{D} a$ and $\eta_n \xrightarrow{D} \eta$ when $n \to \infty$, where $a \in \mathbb{R}$ and η is a random variable, then*

$$\xi_n + \eta_n \xrightarrow{D} a + \eta, \quad \text{as } n \to \infty.$$

Show that if a is a general random variable, these two relations do not hold (construct a counterexample).

3°. *If* $\xi_n \xrightarrow{D} a$ *and* $\eta_n \xrightarrow{D} \eta$ *when* $n \to \infty$*, where* $a \in \mathbb{R}$ *and* η *is a random variable, then*

$$\xi_n \eta_n \xrightarrow{D} a\eta, \quad \text{as } n \to \infty.$$

Would this result continue to hold if we suppose that a is a general random variable?

Proof. 1°. (Anatolii's proof) By the definition of convergence in probability, we want to show that $\lim_{n\to\infty} \mathbb{P}(|\xi_n - a| \geq \varepsilon) = 0$. Let us consider two continuous functions, defined as follows:

$$f_\varepsilon(x) = \begin{cases} 1, & x \in \left(-\infty, a - \frac{3\varepsilon}{2}\right] \bigcup \left[a + \frac{3\varepsilon}{2}, \infty\right) \\ 0, & x \in [a-\varepsilon, a+\varepsilon], \end{cases}$$

$$g_\varepsilon(x) = \begin{cases} 1, & x \in (-\infty, a-\varepsilon] \bigcup [a+\varepsilon, \infty) \\ 0, & x \in \left[a - \frac{\varepsilon}{2}, a + \frac{\varepsilon}{2}\right], \end{cases}$$

Functions f and g are "smoothed" versions of an indicator function, which check if x contains outside of balls of radius ε and $\frac{\varepsilon}{2}$ respectively. Then,

$$\int_{-\infty}^{\infty} f_\varepsilon(x) dF_n(x) \leq \mathbb{P}(|\xi_n - a| \geq \varepsilon) \leq \int_{-\infty}^{\infty} g_\varepsilon(x) dF_n(x),$$

where F_n is a cumulative distribution function of ξ_n. Both left and right sides converge to 0 as $n \to \infty$, since $f_\varepsilon(a) = g_\varepsilon(a) = 0$. Consequently, $\lim_{n\to\infty} \mathbb{P}(|\xi_n - a| \geq \varepsilon) = 0$. It gives the result.

1°. (Alternative proof) By the definition of convergence in probability, we want to show that $\lim_{n\to\infty} \mathbb{P}(|\xi_n - a| \geq \varepsilon) = 0$. Fix some $\varepsilon > 0$. Denote by $B_\varepsilon(a)$ be the open ball of radius ε around point a, and $\bar{B}_\varepsilon(a)$ its complement. Then

$$\mathbb{P}(|\xi_n - a| \geq \varepsilon) = \mathbb{P}\left(\xi_n \in \bar{B}_\varepsilon(a)\right)$$

Then we can observe that:

$$\lim_{n\to\infty} \mathbb{P}(|\xi_n - a| \geq \varepsilon) \leq \limsup_{n\to\infty} \mathbb{P}(|\xi_n - a| \geq \varepsilon) = \limsup_{n\to\infty} \mathbb{P}\left(\xi_n \in \bar{B}_\varepsilon(a)\right) = \mathbb{P}\left(a \in \bar{B}_\varepsilon(a)\right) = 0$$

By definition, it means that the sequence converges to a in probability.

2°*a*. It is a direct consequence of 1°. What we need to show is the following:

$$\mathbb{P}\left(\xi_n + \eta_n \leq t\right) \longrightarrow \mathbb{P}\left(a + \eta \leq t\right) = \mathbb{P}\left(\eta \leq t - a\right).$$

Consider the following event:

$$\{\xi_n + \eta_n - a \le t - a\} = \\ \{\xi_n + \eta_n - a \le t - a, |\xi_n - a| \le \varepsilon\} \bigcup \{\xi_n + \eta_n - a \le t - a, |\xi_n - a| > \varepsilon\}$$

Note that the probability of the second event tends to 0 as $n \to \infty$ (because of the result from 1°). Then, consider

$$\{\eta_n \le t - a - \varepsilon\} \subseteq \{\xi_n + \eta_n - a \le t - a, |\xi_n - a| \le \varepsilon\} \subseteq \{\eta_n \le t - a + \varepsilon\}$$

As a consequence,

$$\mathbb{P}\left(\eta_n \le t - a + \varepsilon\right) \longrightarrow \mathbb{P}\left(\eta \le t - a + \varepsilon\right) \to \mathbb{P}\left(\eta \le t - a\right) \text{ as } \varepsilon \to 0.$$

$2°b$. **(Counterexample)** Consider a sequence ξ_n of Bernoulli variables, such that $\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = 0) = \frac{1}{2}$, and consider $\eta_n = 1 - \xi_n$. It is easy to see that $\xi_n \xrightarrow{D} \xi$ and $\eta_n \xrightarrow{D} \eta$, where ξ and η are again Bernoulli variables taking values 0 and 1 with equal probability. Obviously, $\xi_n + \eta_n = 1$ is not converging in law to the variable $\xi + \eta$, which is taking values $0, 1, 2$ with probabilities $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ respectively.

3°. First note that

$$\xi_n \eta_n = (\xi_n - a)\eta_n + a\eta_n.$$

Note that $a\eta_n \xrightarrow{D} a\eta$ because $\forall a > 0, x \in \mathbb{R}, \quad \mathbb{P}\left(a\eta_n \le x\right) = \mathbb{P}\left(\eta_n \le \frac{x}{a}\right) \to \mathbb{P}\left(\eta \le \frac{x}{a}\right) = \mathbb{P}\left(a\eta \le x\right)$. Also, $\forall C < \infty$

$$\{|\eta_n(\xi_n - a)| > \varepsilon\} \subseteq \{|\eta_n| > C\} \bigcup \left\{|\xi_n - a| > \frac{\varepsilon}{C}\right\},$$

thus

$$\mathbb{P}\left(|\eta_n(\xi_n - a)| > \varepsilon\right) \le \mathbb{P}\left(|\eta_n| > C\right) + \mathbb{P}\left(|\xi_n - a| > \frac{\varepsilon}{C}\right).$$

Note that $\mathbb{P}\left(|\xi_n - a| > \frac{\varepsilon}{C}\right)$ converges to 0 as $n \to \infty$ due to 1°. Now it only remains to note that $\mathbb{P}\left(|\eta_n| > C\right) \to \mathbb{P}\left(|\eta| > C\right) < \frac{\delta}{4}$ for $C = C(\delta)$ sufficiently large. □

Exercise 1.9. *Let $\xi_1, \ldots, \xi_n$ be independent r.v. and let*

$$\xi_{\min} = \min(\xi_1, \ldots, \xi_n), \quad \xi_{\max} = \max(\xi_1, \ldots, \xi_n).$$

1. Show that

$$\mathbb{P}(\xi_{\min} \ge x) = \prod_{i=1}^{n} \mathbb{P}(\xi_i \ge x), \quad \mathbb{P}(\xi_{\max} < x) = \prod_{i=1}^{n} \mathbb{P}(\xi_i < x)$$

2. *Suppose, furthermore, that* $\xi_1, \ldots, \xi_n$ *are identically distributed with uniform distribution* $\mathcal{U}[0,a]$. *Compute* $\mathbb{E}[\xi_{\min}], \mathbb{E}[\xi_{\max}], Var[\xi_{\min}], Var[\xi_{\max}]$

Proof. We consider the variable $\xi_{\max}$. The proof for $\xi_{\min}$ is identical.

1. Thanks to the independence of $\xi_1, \ldots, \xi_n$ we have:

$$\mathbb{P}\left(\max_{i=1,\ldots,n} \xi_i\right) = \mathbb{P}(\xi_1 < x, \ldots, \xi_n < x) = \mathbb{P}(\xi_1 < x) \ldots \mathbb{P}(\xi_n < x).$$

2. Since $\xi_i \sim \mathcal{U}[0,a]$, we have the following c.d.f. $F^*(x)$ of $\xi_{\max}$:

$$F^*(x) = \prod_{i=1}^{n} \frac{x}{a} = \left(\frac{x}{a}\right)^n \quad \forall 0 \le x \le a.$$

From that, we can easily derive the density of $\xi_{\max}$:

$$f^*(x) = \frac{nx^{n-1}}{a^n}.$$

Then, it is easy to obtain the expressions for the first and the second moments:

$$\mathbb{E}[\xi_{\max}] = \int_0^a xn\frac{x^{n-1}}{a^n}dx = \frac{an}{n+1}$$
$$\mathbb{E}[\xi_{\max}^2] = \int_0^a x^2 n\frac{x^{n-1}}{a^n}dx = \frac{a^2 n}{n+2}.$$

From that we compute the varaince:

$$Var[\xi_{\max}] = \mathbb{E}[\xi_{\max}^2] - (\mathbb{E}[\xi_{\max}])^2 = a^2\frac{n}{(n+1)^2(n+2)}.$$

It gives the statement.

□

Exercise 1.10. *Let* $\xi_1, \ldots, \xi_n$ *be i.i.d. Bernoulli r.v. with*

$$\mathbb{P}(\xi_i = 0) = 1 - \lambda_i \Delta, \quad \mathbb{P}(\xi_i = 1) = \lambda_i \Delta,$$

where $\lambda_i > 0$ *and* $\Delta > 0$ *is small. Show that*

$$\mathbb{P}\left(\sum_{i=1}^{n} \xi_i = 1\right) = \left(\sum_{i=1}^{n} \lambda_i\right)\Delta + O(\Delta^2), \quad \mathbb{P}\left(\sum_{i=1}^{n} \xi_i > 1\right) = O(\Delta^2).$$

Proof. Note that

$$\{\xi_1 + \cdots + \xi_n = 1\} = \bigcup_{i=1}^{n} \{\xi_i = 1, \xi_{j \neq i} = 0\}.$$

Since all the variables are independent, the following holds:

$$\begin{aligned}\mathbb{P}(\xi_1 + \cdots + \xi_n = 1) &= \sum_{i=1}^{n} \mathbb{P}\left(\xi_i = 1, \xi_{j\neq i} = 0\right) \\ &= \sum_{i=1}^{n} \mathbb{P}(\xi_i = 1) \prod_{j \neq i} \mathbb{P}\left(\xi_j = 0\right) = \sum_{i=1}^{n} \lambda_i \Delta \prod_{i \neq j} (1 - \lambda_j \Delta) \\ &= \sum_{i=1}^{n} \lambda_i \Delta + O(\Delta^2).\end{aligned}$$

What about the second statement, note that

$$\mathbb{P}\left(\sum_{i=1}^{n} \xi_i > 1\right) = 1 - \mathbb{P}\left(\sum_{i=1}^{n} \xi_i = 0\right) - \mathbb{P}\left(\sum_{i=1}^{n} \xi_i = 1\right)$$

Let us compute the second term:

$$\mathbb{P}\left(\sum_{i=1}^{n} \xi_i = 0\right) = \prod_{i=1}^{n} \mathbb{P}(\xi_i = 0) = \prod_{i=1}^{n} \mathbb{P}(1 - \lambda_i \Delta) = 1 - \sum_{i=1}^{n} \lambda_i \Delta + O(\Delta^2).$$

That, in addition with the result for $\mathbb{P}(\xi_1 + \cdots + \xi_n = 1)$, gives the statement of the exercise. □

Exercise 1.11. 1. *Prove that* $\inf_{a \in \mathbb{R}} \mathbb{E}[(\xi - a)^2]$ *is attained for* $a = \mathbb{E}[\xi]$ *and so*

$$\inf_{a \in \mathbb{R}} \mathbb{E}[(\xi - a)^2] = Var[\xi].$$

2. *Let* ξ *be a nonnegative r.v. with c.d.f.* F *and finite expectation. Prove that*

$$\mathbb{E}[\xi] = \int_0^{\infty} (1 - F(x)) dx.$$

3. *Show, using the result from 2. that if* M *is the median of the c.d.f.* F *of* ξ,

$$\inf_{a \in \mathbb{R}} \mathbb{E}[|\xi - a|] = \mathbb{E}[|\xi - M|].$$

Proof. 1. Trivial (write an expression as a polynom depending on a, take the derivative w.r.t a, find zeroes).

2. Note that by the statement of the exercise, we have

$$\int_t^\infty x dF(x) \to 0 \quad t \to \infty.$$

As $\int_t^\infty x dF(x) \geq t(1-F(t))$, it implies that $t(1-F(t)) \to 0, t \to \infty$. Now we can use the integration by part formula, which results in:

$$\begin{aligned}\mathbb{E}[x] &= \int_0^\infty x dF(x) = -\int_0^\infty x d(1-F(x)) \\ &= -x((1-F(x))\Big|_0^\infty + \int_0^\infty (1-F(x))dx = \int_0^\infty (1-F(x))dx\end{aligned}$$

3. The previous formula actually gives the remaining result. First, we note that $\mathbb{P}(|\xi - a| > x) = \mathbb{P}(\xi > x + a) + \mathbb{P}(\xi > -x + a)$, thus

$$\begin{aligned}\mathbb{E}(|\xi - a|) = \int_0^\infty \mathbb{P}(|\xi - a| > x)\, dx &= \int_0^\infty \mathbb{P}(\xi > x+a)dx + \int_0^\infty \mathbb{P}(\xi > -x + a)\, dx \\ &= \int_a^\infty \mathbb{P}(\xi > z)dz - \int_a^\infty \mathbb{P}(\xi < z)dz\end{aligned}$$

The result can be obtained by computing the derivative w.r.t. a.

□

Exercise 1.12. *Let X_1 and X_2 be two independent r.v. with exponential distribution $\mathscr{E}(\lambda)$. Show that $\min(X_1, X_2)$ and $|X_1 - X_2|$ are r.v. with distributions, respectively, $\mathscr{E}(2\lambda)$ and $\mathscr{E}(\lambda)$.*

Proof. The first result is the direct consequence of Exercise 10. For the second result, consider a r.v. $\zeta = X_1 - X_2$. As both variables X_1 and X_2 are independent, we can use the Fubini theorem, and find the c.d.f of ζ as follows:

$$\begin{aligned}F_\zeta(z) = \mathbb{P}(\zeta < z) &= \int_{x\geq 0, y \geq 0, x-y\leq z} dF(x)dF(y) = \int_{x,y\geq 0} \mathbb{1}_{x-y\leq z} dF(x)dF(y) \\ &= \int_0^\infty dF(x)\left[\int_0^\infty \mathbb{1}_{y\geq x-z} dF(y)\right] \\ &= \int_0^\infty dF(x)\left[\mathbb{1}_{x-z\geq 0}\int_{x-z}^\infty dF(y) + \mathbb{1}_{x-z<0}\int_0^\infty dF(y)\right] \\ &= \int_0^\infty dF(x)\left[\mathbb{1}_{x\geq z}(1-F(x-z)) + \mathbb{1}_{x<z}\right]\end{aligned}$$

Then, two cases are possible:

$z < 0$:

$$F_\zeta(z) = \int_0^\infty dF(x)(1-F(x-z)) = e^{\lambda z}\lambda \int_0^\infty e^{-2\lambda x}dx = \frac{1}{2}e^{\lambda z}$$

$z \geq 0$:

$$F_\zeta(z) = \int_0^z dF(x) + \int_z^\infty dF(x)(1 - F(x-z)) = F(z) + \lambda \int_z^\infty e^{\lambda(z-x)} e^{-\lambda x} dx$$
$$= (1 - e^{-\lambda z}) + \frac{1}{2} e^{-\lambda z} = 1 - \frac{e^{-\lambda z}}{2}$$

It only remains to note that $F_{|\zeta|}(x) = F_\zeta(x) - F_\zeta(-x)$ for all $x \geq 0$. □

Exercise 1.14. *Suppose that r.v. $\xi_1, \dots, \xi_n$ are mutually independent and identically distributed with the c.d.f. F. For $x \in \mathbb{R}$, let us define the random variable $\hat{F}_n(x) = \frac{1}{n}\mu_n$, where μ_n is the number of $\xi_1, \dots, \xi_n$ which satisfy $\xi_k \leq x$. Show that for any x,*

$$\hat{F}_n(x) \xrightarrow{P} F(x).$$

The function $\hat{F}_n(x)$ is called ***the empirical distribution function****.*

Proof. Consider a sequence of random variables $\zeta_1, \dots, \zeta_n$ such that $\zeta_i = \mathbb{1}_{\xi_k \leq x}$. Note that $\{\zeta_i\}_{i=1,n}$ is a sequence of i.i.d. Bernoulli random varaibles with the probability of success $F(x)$. Observe that

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \zeta_i.$$

$n\hat{F}_n(x)$ is a Binomial random variable with the expectation and variance being $F(x)$ and $\frac{F(x)(1-F(x))}{n}$ respectively. Then, by Chebyshev's inequality, we have the following result $\forall \varepsilon > 0$:

$$\mathbb{P}\left(|F_n(x) - F(x)| \geq \varepsilon\right) \leq \frac{F(x)(1 - F(x))}{n\varepsilon^2}$$

The right part converges to 0 as $n \to \infty$, which gives the result. □