# Tinbergen Institute
# Measure Theory and Asymptotic Statistics
# Exam Questions

---

1. If $X$ and $Y$ are *independent* random variables with $\mathbb{E}\,|X| < \infty$ and $\mathbb{E}\,|Y| < \infty$ (assumed to hold throughout this exercise), then the product formula $\mathbb{E}\,(XY) = \mathbb{E}\,X \cdot \mathbb{E}\,Y$ holds. To show this you have to apply (parts of) the standard machine[1] a couple of times.

   (a) First a special case. Let $X$ be positive but arbitrary otherwise, and $Y = \mathbf{1}_A$ for some set $A \in \mathcal{F}$. Use the standard machine to show that $\mathbb{E}\,(X\mathbf{1}_A) = \mathbb{E}\,X \cdot \mathbb{P}(A)$.

   (b) Prove now, using the previous item and the standard machine again, the product formula for $X \geq 0$ and $Y \geq 0$.

   (c) Why are $X^+$ and $Y^-$ also independent random variables?

   (d) Complete the proof for arbitrary $X$ and $Y$.

2. Let $X$ and $Y$ be random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} = \sigma(Y)$.

   (a) Show that the collection of events $\{Y \in B\}$, where $B$ runs through the Borel sets $\mathcal{B}(\mathbb{R})$, forms a $\sigma$-algebra (so you show that it has all the defining properties of a $\sigma$-algebra). This $\sigma$-algebra will be denoted $\mathcal{H}$.

   (b) Show the two inclusions $\mathcal{H} \subset \mathcal{G}$ and $\mathcal{G} \subset \mathcal{H}$. For the latter you need the 'minimality property' of $\sigma(Y)$.

   (c) Let $X = \mathbf{1}_G$ for some $G \in \mathcal{G}$. Find a function $f : \mathbb{R} \to [0, 1]$ that is Borel-measurable (and check this property!) such that $X = f(Y)$.

   (d) Use the standard machine to prove the following result. If $X$ is $\mathcal{G}$-measurable, then there exists a Borel-measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $X = f(Y)$.

3. Let $X_1, X_2, \ldots$ be random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that the $X_i$ are nonnegative and let $S_n = \sum_{i=1}^n X_i$ for $n \geq 1$. It is known that the $S_n$ are random variables (measurable functions) as well. We define $S(\omega) = \lim_{n\to\infty} S_n(\omega)$, which exists for every $\omega \in \Omega$ but may be infinite.

---

[1] Recall that the standard machine is a method of proving along steps: (1) for indicator functions; (2) for nonnegative simple functions; (3) for nonnegative functions by approximation with simple functions (the approximating sequence always exists); (4) general case.

(a) Show that $S$ is a random variable (*Hint:* show first that $\{S > a\} = \bigcup_{n=1}^{\infty}\{S_n > a\}$ for $a > 0$).

(b) Note that $\mathbb{E}\, S \leq \infty$ is well defined. Show that $\mathbb{E}\, S = \sum_{i=1}^{\infty} \mathbb{E}\, X_i$.

(c) Assume that $\sum_{i=1}^{\infty} \mathbb{E}\, X_i < \infty$. Show that $\mathbb{P}(S < \infty) = 1$.

4. Let $X$ be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A well known property is that $\mathbb{E}\, X = 0$ if $X = 0$ a.s. In this exercise you will show this.

   (a) Suppose that $X$ assumes finitely many values $y_0, y_1, \ldots, y_n$ and also that $X = 0$ a.s. Show that $\mathbb{E}\, X = 0$.

   (b) Suppose that $X \geq 0$, but also $X = 0$ a.s. Argue by using lower Lebesgue sums and the previous item that $\mathbb{E}\, X = 0$.

   (c) Let $X$ be arbitrary but still $X = 0$ a.s. Show again that $\mathbb{E}\, X = 0$.

5. Recall the definition of *infimum*, written as inf. If $x_1, x_2, \ldots$ is a finite or infinite sequence of real numbers, then $x = \inf\{x_1, x_2, \ldots\}$ iff (1) $x \leq x_k$ for all $k$ and (2) if $y > x$, there exists $x_k$ such that $x_k < y$. It may happen that $x = -\infty$. For finite sequences $x_1, \ldots, x_n$, $\inf\{x_1, \ldots, x_n\}$ is the *minimum* of the $x_k$. An example with an infinite sequence is $\inf\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\} = 0$, another example is $\inf\{1, \frac{1}{2}, 1, \frac{1}{3}, 1, \frac{1}{4}, \ldots\} = 0$.

   If we have an infinite sequence of random variables $X_1, X_2, \ldots$, we say that the random variable $X$ is $\inf\{X_1, X_2, \ldots\}$ if for every $\omega \in \Omega$ one has $X(\omega) = \inf\{X_1(\omega), X_2(\omega), \ldots\}$. From now on we assume to have a sequence of *nonnegative* random variables $X_1, X_2, \ldots$. For each $n$ we define the random variable $Y_n := \inf\{X_n, X_{n+1}, X_{n+2}, \ldots\}$, also written as $Y_n = \inf_{m \geq n} X_m$.

   (a) Show that (each) $Y_n$ is a random variable by considering events like $\{Y_n \geq a\}$.

   (b) Show that the $Y_n$ form an *increasing* sequence of random variables. They then have a limit $Y_\infty \leq \infty$.

   (c) Show that $Y_n \leq X_m$ for all $m \geq n$, and conclude that $\mathbb{E}\, Y_n \leq y_n := \inf\{\mathbb{E}\, X_n, \mathbb{E}\, X_{n+1}, \ldots\}$. Note that the $y_n$ form an increasing sequence too.

   (d) Show that $\mathbb{E}\, Y_\infty \leq \lim_{n\to\infty} y_n$. *This property is often written as* $\mathbb{E} \lim_{n\to\infty} \inf_{m\geq n} X_m \leq \lim_{n\to\infty} \inf_{m\geq n} \mathbb{E}\, X_m$, *and is known as Fatou's lemma.*

(e) In the previous item, a strict inequality may occur. Consider thereto the probability space with $\Omega = (0, 1)$, $\mathcal{F}$ the Borel sets in $(0, 1)$ and $\mathbb{P}$ the Lebesgue measure. We take $X_n(\omega) = n\,\mathbf{1}_{(0,1/n)}(\omega)$. Show that indeed strict inequality now takes place in Fatou's lemma (so you compute both sides of the inequality).

6. In this exercise we need limits of sequences of subsets of a given set $\Omega$, which we define in two cases. Suppose that we have an *increasing* sequence of sets $A_n$ $(n \geq 0)$, i.e. $A_n \subset A_{n+1}$ for all $n \geq 0$. Then we define $\overline{A} = \lim_{n\to\infty} A_n := \bigcup_{n=0}^{\infty} A_n$. If the sequence is *decreasing*, $A_n \supset A_{n+1}$ for all $n$, we define $\underline{A} = \lim_{n\to\infty} A_n := \bigcap_{n=0}^{\infty} A_n$. We work with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we consider an increasing sequence of events $A_n$ (so $A_n \in \mathcal{F}$ for all $n$). Let $D_0 = A_0$ and $D_n = A_n \setminus A_{n-1}$ for $n \geq 1$.

   (a) Show that $\mathbb{P}(A_n) = \sum_{k=0}^{n} \mathbb{P}(D_k)$.

   (b) Show that $\overline{A} = \bigcup_{k=0}^{\infty} D_k$.

   (c) Show that $\mathbb{P}(A_n) \to \mathbb{P}(\overline{A})$ for $n \to \infty$.

   (d) Suppose that events $B_n$ $(n \geq 0)$ form a decreasing sequence. Show that $\mathbb{P}(B_n) \to \mathbb{P}(\underline{B})$. (Hint: consider the $B_n^c$.)

7. Let $X, Y$ be random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, so they are both $\mathcal{F}$-measurable.

   (a) Let $c \in \mathbb{R}$. Show (make a sketch!) that $\{(x, y) \in \mathbb{R}^2 : x + y > c\} = \bigcup_{q\in\mathbb{Q}} \{(x, y) \in \mathbb{R}^2 : x > q, y > c - q\}$.

   (b) Show that $X+Y$ is also $\mathcal{F}$-measurable. NB: For this it is sufficient to show that $\{X + Y > c\} \in \mathcal{F}$ for all $c \in \mathbb{R}$.

8. Let $x_1, x_2, \ldots$ be a sequence of real numbers. We put, for $n \geq 1$, $\overline{x}_n = \sup\{x_n, x_{n+1}, \ldots\}$ and $\underline{x}_n = \inf\{x_n, x_{n+1}, \ldots\}$. Note that the $\overline{x}_n$ form a decreasing sequence and the $\underline{x}_n$ an increasing one, and hence both sequences have a limit, denoted $\overline{x}$ and $\underline{x}$ respectively. One always has $\overline{x} \geq \underline{x}$ and $\overline{x} = \inf\{\overline{x}_1, \overline{x}_2, \ldots\}$. Moreover, the original sequence with the $x_n$ has a limit $x$ iff $x = \overline{x} = \underline{x}$.

   Consider now a sequence of random variables $X_n$ defined on some $(\Omega, \mathcal{F}, \mathbb{P})$. As these are measurable functions, we can define $\overline{X}_n$ as the function s.t. $\overline{X}_n(\omega) = \sup\{X_n(\omega), X_{n+1}(\omega), \ldots\}$ and likewise $\underline{X}_n$, $\overline{X}$, $\underline{X}$.

(a) Consider $E_a = \{\overline{X}_n \leq a\}$ for arbitrary $a \in \mathbb{R}$. Show that $E_a \in \mathcal{F}$ and conclude that $\overline{X}_n$ is a random variable (for every $n$).

(b) Show that $\underline{X}_n$ is a random variable.

(c) Show that $\overline{X}$ and $\underline{X}$ are random variables too.

(d) Show that $\{\omega : \lim_{n\to\infty} X_n(\omega) \text{ exists}\} = \{\omega : \overline{X}(\omega) - \underline{X}(\omega) \leq 0\}$ and that this set belongs to $\mathcal{F}$.

(e) Assume that $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ exists for every $\omega$. Show that $X$ is a random variable.

9. Consider a sequence of random variables $X_n$ defined on some $(\Omega, \mathcal{F}, \mathbb{P})$ and put $S_n = \sum_{k=1}^n X_k$ for $n \geq 1$.

   (a) Assume all $X_n \geq 0$. Show that $\mathbb{E} \sum_{k=1}^\infty X_k = \sum_{k=1}^\infty \mathbb{E}\, X_k$. *Hint: apply the Monotone Convergence Theorem to the* $S_n$.

   From here on the assumption that the $X_n$ are nonnegative is dropped.

   (b) Show that $\mathbb{E} \sum_{k=1}^\infty |X_k| = \sum_{k=1}^\infty \mathbb{E}\, |X_k|$.

   (c) Assume $\sum_{k=1}^\infty \mathbb{E}\, |X_k| < \infty$. Show that $\mathbb{E} \sum_{k=1}^\infty X_k = \sum_{k=1}^\infty \mathbb{E}\, X_k$.

10. Consider a probability space and a sequence of events $(E_n)_{n\geq 1}$. The event $E := \limsup E_n$ is defined as $E = \bigcap_{n=1}^\infty U_n$, where $U_n = \bigcup_{m=n}^\infty E_m$. Note that the $U_n$ form a decreasing sequence. Further we have $E^c = \bigcup_{n=1}^\infty D_n$, with $D_n = \bigcap_{m=n}^\infty E_m^c$. We also write $D_n^N = \bigcap_{m=n}^N E_m^c$ for $N \geq n$.

    (a) Show that $\mathbb{P}(E) \leq \mathbb{P}(U_n)$ for every $n$ and that $\mathbb{P}(E) = 0$ if $\sum_{n=1}^\infty \mathbb{P}(E_n) < \infty$.

    From now on we assume that the $E_n$ are independent events.

    (b) Show that $\mathbb{P}(D_n^N) \leq \exp(-\sum_{m=n}^N \mathbb{P}(E_m))$. [Recall $e^{-x} \geq 1 - x$.]

    (c) Assume further also that $\sum_{n=1}^\infty \mathbb{P}(E_n) = \infty$. Show that $\mathbb{P}(D_n) = 0$ for all $n$ and deduce that $\mathbb{P}(E) = 1$.

    The conclusions in (a) and (c) are together known as the *Borel-Cantelli lemma.*

11. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a random variable $X$ is defined with $\mathbb{E}\, |X| < \infty$. Let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$ and let $\hat{X} = \mathbb{E}\, [X|\mathcal{G}]$.

(a) Show that $\mathbb{E}\,\hat{X}^+ \leq \mathbb{E}\,|X|\mathbf{1}_{\{\hat{X}>0\}}$. *Hint: use* $x^+ = x\mathbf{1}_{\{x>0\}}$ *and the definition of conditional expectation.*

(b) Show that $\mathbb{E}\,|\hat{X}| \leq \mathbb{E}\,|X|$.

12. Consider the Pareto distribution with parameters $\alpha, \mu > 0$. This distribution has density

$$p_{\alpha,\mu}(x) = \frac{\alpha\mu^\alpha}{x^{\alpha+1}}\mathbf{1}_{\{x\geq\mu\}}.$$

Let $Y_1, \ldots, Y_n$ be a sample from this distribution, and $X_i = \log Y_i$, $i = 1, \ldots, n$. It is possible to show that $\mathbb{E}\,X_1 = \log\mu + \frac{1}{\alpha}$ and $\operatorname{Var} X_1 = \frac{1}{\alpha^2}$. Suppose $\mu$ is known.

(a) Let $\hat{\alpha}_n$ be the maximum likelihood estimator of $\alpha$. Show that $\hat{\alpha}_n = \frac{1}{\overline{X}_n - \log\mu}$.

(b) Deduce from the Central limit theorem for averages and the Delta method that $\sqrt{n}(\hat{\alpha}_n - \alpha)$ converges in distribution to $N(0, \alpha^2)$.

In the sequel also $\mu$ is unknown.

(c) Show that the maximum likelihood estimator of $\mu$ is $\hat{\mu}_n = \exp(\underline{X}_n)$, where $\underline{X}_n = \min\{X_1, \ldots, X_n\}$.

(d) Show that $X_i - \log\mu$ has an exponential distribution and that $\mathbb{P}(n(\underline{X}_n - \log\mu) > c) = \exp(-c\alpha)$ for any $c > 0$.

(e) Show that $\mathbb{P}(n(\hat{\mu}_n - \mu) > c) \to \exp(-c\alpha/\mu)$ for any $c > 0$. [Depending on the method, you may need $\log(1 + x) = x + O(x^2)$ for $x \to 0$.]

(f) What is the (obvious) maximum likelihood estimator, call it $\hat{\alpha}_n$ again, of $\alpha$ in the present situation? Argue that the limit distribution of $\sqrt{n}(\hat{\alpha}_n - \alpha)$ is the same as in question **??**.

13. Let $X_1, \ldots, X_n$ be independent random variables with a $N(\theta, \theta^2)$ distribution. Here $\theta \neq 0$ is an arbitrary real parameter. We consider the maximum likelihood estimator $\hat{\theta}_n$, a maximizer of $M_n(\theta)$, where $M_n(\theta) = \frac{1}{n}\sum_{i=1}^n \log\frac{p_\theta(X_i)}{p_{\theta_0}(X_i)}$ and $p_\theta(X_i)$ the likelihood of $\theta$ when $X_i$ is observed and $\theta_0 \neq 0$ is the true parameter. Probabilities or expectations below are taken under the true parameter. It turns out that $\hat{\theta}_n = -\frac{1}{2}\overline{X}_n + \operatorname{sign}(\overline{X}_n)\sqrt{\frac{1}{4}\overline{X}_n^2 + \overline{X_n^2}}$. Here $\overline{X}_n$ is the average of the $X_i$ and $\overline{X_n^2}$ is the average of the $X_i^2$. In the computations you may need the following results: $\mathbb{E}_{\theta_0}X^3 = 4\theta_0^3$, $\mathbb{E}_{\theta_0}X^4 = 10\theta_0^4$.

(a) Show that $M_n(\theta) = -\frac{1}{2}\log\frac{\theta^2}{\theta_0^2} - \frac{1}{2n}\sum_{i=1}^n(\frac{X_i}{\theta}-1)^2 + \frac{1}{2n}\sum_{i=1}^n(\frac{X_i}{\theta_0}-1)^2$.

(b) Show: $M_n(\theta) \xrightarrow{\mathrm{P}} M(\theta) = -\frac{1}{2}\log\theta^2 + \frac{1}{2}\log\theta_0^2 - \frac{1}{2}\frac{\theta_0^2}{\theta^2} - \frac{1}{2}(\frac{\theta_0}{\theta}-1)^2 + \frac{1}{2}$.

(c) Show that $M(\theta_0) = 0$ and that $\theta_0$ is the maximizer of $M$.

(d) We expect that $\hat{\theta}_n$ is consistent. Show this by a direct argument, using the law of large numbers for $\overline{X}_n$ and $\overline{X_n^2}$.

(e) Let $\Psi_n(\theta) = \dot{M}_n(\theta)$ and $\Psi(\theta) = \dot{M}(\theta)$. What would you expect (ignoring certain conditions) for the asymptotic variance of $\hat{\theta}_n$?

(f) Show that the Fisher information $I_{\theta_0}$ equals $3\theta_0^{-2}$.

(g) The central limit theorem gives $\sqrt{n}\begin{pmatrix}\overline{X}_n - \theta_0 \\ \overline{X_n^2} - 2\theta_0^2\end{pmatrix} \rightsquigarrow N_2(0, \Sigma(\theta_0))$, where $\Sigma(\theta_0) = \begin{pmatrix}\theta_0^2 & 2\theta_0^3 \\ 2\theta_0^3 & 6\theta_0^4\end{pmatrix}$. Use this and the fact that $\hat{\theta}_n = \phi(\overline{X}_n, \overline{X_n^2})$ (for which $\phi$?) to deduce that indeed $\hat{\theta}_n$ is asymptotically normal with variance given by the inverse of the Fisher information.

14. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a random variable $X$ that has a standard exponential distribution, $\mathbb{P}(X \leq x) = 1 - e^{-x}$ for $x \geq 0$. Let $\lambda$ be a positive constant and consider $Z = \lambda\exp(-(\lambda-1)X)$, a positive random variable. Using $Z$ we define a new measure $\mathbb{P}'$ on $\mathcal{F}$ by $\mathbb{P}'(F) = \mathbb{E}\,[1_F Z]$ (theory guarantees that $\mathbb{P}'$ is indeed a measure).

(a) Show that $\mathbb{E}\,Z = 1$. Is $\mathbb{P}'$ a probability measure?

(b) Show (by computing an integral) that $\mathbb{P}'(X \leq x) = 1 - e^{-\lambda x}$. [It follows that $X$ has an exponential distribution with parameter $\lambda$ under $\mathbb{P}'$.]

15. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $X$ be a nonnegative random variable defined on it. Let $h$ be a monotone increasing function, $h : [0, \infty) \to [0, \infty)$ with $h(0) = 0$. We will need the product space $S = \Omega \times [0, \infty)$ with the product $\sigma$-algebra $\mathcal{F} \times \mathcal{B}[0, \infty)$ and the product measure $\mathbb{P} \times \lambda$, where $\lambda$ is the Lebesgue measure on $\mathcal{B}[0, \infty)$.

(a) Show that $h$ is Borel-measurable. [*Hint: consider the sets* $\{h \leq c\}$ *for* $c > 0$*; these sets have a nice structure.*]

(b) We can extend $h$ to a function on $S$ by putting $h(\omega, x) = h(x)$. Show (use part (a)) that $h$ is $\mathcal{F} \times \mathcal{B}[0, \infty)$-measurable. In a similar way the identity map on $[0, \infty)$ (i.e. $u \mapsto u$) can be considered $\mathcal{F} \times \mathcal{B}[0, \infty)$-measurable.

(c) Show that it follows that the set $E := \{(\omega, u) \in \Omega \times [0, \infty) : h(X(\omega)) \geq u\}$ is $\mathcal{F} \times \mathcal{B}[0, \infty)$-measurable.

(d) Use the set $E$ above and Fubini's theorem to show that $\mathbb{E}\, h(X) = \int_0^\infty (1 - F(h^{-1}(u)) \, \mathrm{d}u$, where $F$ is the distribution function of $X$ and $h^{-1}$ is the inverse function of $h$.

16. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a random variable $X$. Assume $\mathbb{E}\,|X|$ finite, $\mathcal{G}$ a sub-$\sigma$-algebra of $\mathcal{F}$ and let $\hat{X}$ be (a version of) $\mathbb{E}\,[X|\mathcal{G}]$.

(a) Show that $\hat{X} \leq \mathbb{E}\,[|X|\,|\mathcal{G}]$ and conclude that $|\hat{X}| \leq \mathbb{E}\,[|X|\,|\mathcal{G}]$.

(b) Let $f$ be a convex differentiable function. Then for every $x, x_0$ it holds that $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$. Note that $y = f(x_0) + f'(x_0)(x - x_0)$ gives the tangent line of $f$ at $x_0$. Verify this inequality by a sketch for $f(x) = |x|^2$. Use the inequality with $x = X$ and $x_0 = \hat{X}$, assuming that $\mathbb{E}\,|X|^p < \infty$, to show that $\mathbb{E}\,|X|^p \geq \mathbb{E}\,|\hat{X}|^p$ for $p > 1$.

Consider also a sequence $(\mathcal{G}_n)$ of sub-$\sigma$-algebras of $\mathcal{F}$ and let, for each $n$, $X_n$ be (a version of) the conditional expectation $\mathbb{E}\,[X|\mathcal{G}_n]$.

(c) Suppose that for some $a > 0$ it holds that $\mathbb{E}\,|X|^{1+a} < \infty$. Show that $|X_n|\mathbf{1}_{\{|X_n|>m\}} \leq \frac{|X_n|^{1+a}}{m^a}$ and deduce that $\sup_n \mathbb{E}\,|X_n|\mathbf{1}_{\{|X_n|>m\}} \to 0$ for $m \to \infty$. [A sequence $(X_n)$ with this property is said to be uniformly integrable.]

17. Under certain conditions, among them continuous dependence of $I_\theta$ on $\theta$, one has that $\sqrt{n}(\hat{\theta}_n - \theta_0)$ has an asymptotically normal $N(0, \frac{1}{I_{\theta_0}})$ distribution. Here $\hat{\theta}_n$ is the maximum likelihood estimator, which is assumed to be consistent, based on a sample from a distribution with density $p_{\theta_0}$ and $\theta_0$ is one-dimensional.

(a) Show that $\sqrt{nI_{\hat{\theta}_n}}(\hat{\theta}_n - \theta_0) \rightsquigarrow N(0, 1)$.

Consider a sample $X_1, \ldots, X_n$ from an exponential distribution with density $\frac{1}{\theta}\exp(-x/\theta)$. Later we will use the different parametrization with $\lambda = 1/\theta$. Recall that $\mathbb{E}\, X_1 = \theta_0$ and $\operatorname{Var} X_1 = \theta_0^2$, $\theta_0$ is the 'true' parameter. Consider the maximum likelihood estimator $\hat{\theta}_n$.

(b) Show by using the ordinary central limit theorem that $\sqrt{n}(\hat{\theta}_n - \theta_0) \rightsquigarrow N(0, \theta_0^2)$.

(c) Compute the maximum likelihood estimator $\hat{\lambda}_n$ of $\lambda = 1/\theta_0$ and show by the delta method that $\sqrt{n}(\hat{\lambda}_n - \lambda_0) \rightsquigarrow N(0, \lambda_0^2)$.

(d) Compute (under the alternative parametrization) the Fisher information $I_{\lambda_0}$ and show that the answer of question **??** agrees with the general result on the asymptotic behaviour on maximum likelihood estimators.

(e) Give a confidence interval of level $1 - \alpha$ for $\theta_0$.

18. In the formula for the asymptotic distribution of the Huber estimator one needs the derivative $V_\theta$ w.r.t. $\theta$ of $P\psi_\theta = \int \psi(x - \theta)p(x)\,\mathrm{d}x$, where $\psi$ is the usual Huber function and $p$ a probability density function. We define the measure $\mu$ on $\mathcal{B}(\mathbb{R})$ by $\mu(B) = \int_B \mathbf{1}_{[-k,k]}(x)\,\mathrm{d}x$ (the integral can be seen as a Riemann integral and as an integral w.r.t. the Lebesgue measure $\lambda$). The function $\psi$ and the measure $\mu$ are related by $\psi(x) + k = \mu((-\infty, x]) = \int_{(-\infty,x]} \mathrm{d}\mu$. It follows that $\mu \ll \lambda$ and for a measurable function $h$ for which the integrals exist, one has $\mu(h) = \int \mathbf{1}_{[-k,k]}(x)h(x)\,\mathrm{d}x$.

(a) Understanding that $\psi(x - \theta)$ can be written as an integral minus the constant $k$, show by application of Fubini's theorem that $P\psi_\theta = k - \int_{-k}^{k} F(u + \theta)\,\mathrm{d}u$, where $F$ is the distribution function with density $p$.

(b) Show that $V_\theta = F(\theta - k) - F(\theta + k)$.

(c) As an alternative to the ordinary Huber function, one can also use the scaled Huber function $\bar{\psi}_k = \frac{1}{k}\psi$. Note that $\lim_{k\to 0} \bar{\psi}_k(x) = \operatorname{sign}(x)$. Show that the asymptotic distribution of the Huber estimator doesn't change if we replace $\psi$ by $\bar{\psi}_k$ in the estimation procedure.

(d) Let $\bar{V}_{\theta,k}$ be the derivative of $\int \bar{\psi}_k(x - \theta)p(x)\,\mathrm{d}x$. Compute the limit, you may assume that $p$ is continuous, of $\bar{V}_{\theta,k}$ for $k \to 0$. Why can you expect this result?

19. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $X$ be a nonnegative random variable defined on it. For $t \geq 0$ put $\phi(t) = \mathbb{E}\exp(-tX)$.

(a) Show that $0 \leq \phi(t) \leq 1$ for all $t \geq 0$.

(b) Show (use dominated convergence) that $\phi$ is continuous at any $t \geq 0$,
i.e. $\lim_{h\to 0} \phi(t + h) = \phi(t)$. [*Note: for $t = 0$ one only has right continuity.*]

(c) Assume $\mathbb{E}\, X < \infty$ and consider $r(h) := \frac{1}{h}(\phi(t) - \phi(t+h)) = \frac{1}{h}\mathbb{E}\left(e^{-tX}(1 - e^{-hX})\right)$ for any fixed $t$ and $h$ such that $t + h \geq 0$. Show that $r(h) \to \mathbb{E}\left[e^{-tX}X\right]$ for $h \to 0$. Deduce that $\phi'(t) = -\mathbb{E}\left[e^{-tX}X\right]$. [*You may use that* $|e^{-a} - e^{-b}| \leq |a - b|$ *for* $a, b \geq 0$ *and* $\frac{1}{h}(1 - e^{-hu}) \to u$ *for* $h \to 0$.]

(d) Assume that $\mathbb{E}\, X^2 < \infty$. Knowing what $\phi'(t)$ is, you show that $\phi''(t) = \mathbb{E}\left[e^{-tX}X^2\right]$.

(e) Look at $\phi(0)$, $\phi'(0)$ and $\phi''(0)$. Guess what $\phi^{(k)}(0)$ should be ($k \in \mathbb{N}$), and what the needed assumption should be.

20. Let $f, f_1, f_2, \ldots$ be density functions of probability distributions on $(\mathbb{R}, \mathcal{B})$, they are nonnegative, measurable and their integrals w.r.t. the Lebesgue measure $\lambda$ equal 1.

(a) Show that $(f_n - f)^- \leq f$ and $|f_n - f| = (f_n - f) + 2(f_n - f)^-$.

(b) Assume that $f_n \to f$ a.e. Show that $\int |f_n - f| \, \mathrm{d}\lambda \to 0$.

(c) Assume that $f_n \to f$ a.e. Show that $F_n(x) \to F(x)$ for all $x \in \mathbb{R}$.

21. Assume that $X, X_1, X_2, \ldots$ are $\mathbb{R}^1$-valued random variables. They have the property that $\lim_{n\to\infty} \mathbb{E}\, h(X_n) = \mathbb{E}\, h(X)$ for every bounded and continuous function $h$ on $\mathbb{R}$. For every $x \in \mathbb{R}$ and $m \in \mathbb{N}$ we define $h_{x,m} : \mathbb{R} \to [0, 1]$ by

$$h_{x,m}(u) = \begin{cases} 1 & \text{if } u < x, \\ 1 + m(x - u) & \text{if } x \leq u \leq x + \frac{1}{m}, \\ 0 & \text{if } u > x + \frac{1}{m}. \end{cases}$$

Note that $\mathbf{1}_{(-\infty,x]}(u) \leq h_{x,m}(u) \leq \mathbf{1}_{(-\infty,x+\frac{1}{m}]}(u)$ (draw a picture, if you like).

(a) Show that $\mathbb{P}(X_n \leq x) \leq \mathbb{E}\, h_{x,m}(X_n)$, $\mathbb{E}\, h_{x,m}(X) \leq \mathbb{P}(X \leq x + \frac{1}{m})$, and conclude that $\limsup_{n\to\infty} \mathbb{P}(X_n \leq x) \leq \mathbb{P}(X \leq x + \frac{1}{m})$.

(b) Show that $\liminf_{n\to\infty} \mathbb{P}(X_n \leq x) \geq \mathbb{P}(X \leq x - \frac{1}{m})$.

(c) Show that $X_n \rightsquigarrow X$.

22. Let $\psi$ be the usual Huber function (depending on some $k > 0$),

$$\psi(u) = \begin{cases} -k & \text{if } u < -k, \\ u & \text{if } -k \leq u \leq k, \\ k & \text{if } u > k. \end{cases}$$

We also have a sample $X_1, \dots, X_n$ of IID random variables with a common density function $p_{\theta_0}$ that is everywhere strictly positive. The parameter $\theta_0$ is to be estimated. We consider for $\theta \in \mathbb{R}$ the random variables $\Psi_n(\theta) = \frac{1}{n}\sum_{i=1}^n \psi(X_i - \theta)$ for $\theta \in \mathbb{R}$, and, with $\psi_\theta(x) = \psi(x-\theta)$, $\Psi(\theta) = \mathbb{E}\,\Psi_n(\theta) = \mathbb{E}\,\psi_\theta(X_1) = P\psi_\theta$.

(a) The equation $\Psi_n(\theta) = 0$ has a solution, $\hat\theta_n$ say. Why?

(b) Show by a direct computation of the expectation $\mathbb{E}\,\psi_\theta(X_1)$ (you have to compute an integral) that $\Psi(\theta) = k - \int_{\theta-k}^{\theta+k} F_{\theta_0}(x)\,\mathrm{d}x$, where $F_{\theta_0}$ is the distribution function of $p_{\theta_0}$. [*The integral you can compute as the sum of three integrals, one of them you further compute using integration by parts. Or, you do integration by parts on a single integral.*]

It is now also given that $p_{\theta_0}(x) = p(x-\theta_0)$, where $p$ is a density function that is symmetric around zero.

(c) Show that $\dot\Psi(\theta) < 0$ for every $\theta$ and that $\Psi(\theta) = 0$ iff $\theta = \theta_0$. [*In your answer you may first show that $\Psi(\theta_0)$ is the integral of an odd function; recall that $f$ is odd if $f(-x) = -f(x)$.*]

(d) Argue that $\hat\theta_n$ is a consistent estimator of $\theta_0$.

(e) Show that $\sqrt{n}(\hat\theta_n - \theta_0)$ is asymptotically normal with variance $\sigma^2$ equal to

$$\sigma^2 = \frac{\int_{-k}^k x^2 p(x)\,\mathrm{d}x + k^2\int_{|x|\geq k} p(x)\,\mathrm{d}x}{(\int_{-k}^k p(x)\,\mathrm{d}x)^2}.$$

23. Let $X_1, \dots, X_n$ be a sample from a distribution with a positive and finite variance $\sigma^2$. Independently from this sample there is another sample $Y_1, \dots, Y_{2n}$ from a distribution with positive and finite variance $\tau^2$. Note that in the second case the sample size is twice as big as in the first case. $X_i$ and $Y_i$ are one dimensional. The parameter $\sigma^2$ is estimated by $\hat\sigma^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$ and $\tau^2$ is estimated by $\hat\tau^2 = \frac{1}{2n}\sum_{i=1}^{2n}(Y_i - \overline{Y})^2$. It is known that $\sqrt{n}(\hat\sigma^2 - \sigma^2)$ has a limit law which is normal with variance equal to $\kappa_X - \sigma^4$, where $\kappa_X = \mathbb{E}\,(X_1 - \mathbb{E}\,X_1)^4$, which is assumed to be finite (in what follows only the constant $\kappa_X$ itself matters). Of course there is a parallel result for $\hat\tau^2$ (but note again the different sample size there).

(a) What is the limit law of the random vector $\sqrt{n}\begin{pmatrix}\hat\sigma^2 - \sigma^2 \\ \hat\tau^2 - \tau^2\end{pmatrix}$?

(b) We are interested in estimating the ratio $r = \sigma^2/\tau^2$ which we do by $\hat{r} = \hat{\sigma}^2/\hat{\tau}^2$. What is the limit law of $\sqrt{n}(\hat{r} - r)$? [*There is a certain method to apply here.*]

(c) If the distribution of the $X_i$ is normal, then it is known that $\kappa_X = 3\sigma^4$ and a similar result holds for normal $Y_i$. Show that in this case the limit variance of $\sqrt{n}(\hat{r} - r)$ is equal to $3\sigma^4/\tau^4$.

(d) Give a $(1-\alpha)$-confidence interval for $r$ under the normality assumptions of the previous item.

24. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $E_1, E_2, \ldots$ be an arbitrary sequence of events, put $U_n = \bigcup_{m \geq n} E_m$, $n \geq 1$.

(a) Write the limit $U$ of the $U_n$ in terms of the $E_n$.

(b) Show the following. If $\sum_{n \geq 1} \mathbb{P}(E_n) < \infty$, then $\mathbb{P}(\limsup E_n) = 0$. [*Hint: Find upper and lower bounds of* $\mathbb{P}(U_n)$.]

It is further assumed that the $E_n$ are *independent* events (then also their complements $E_n^c$ are independent) and $\sum_{n \geq 1} \mathbb{P}(E_n) = \infty$. Put $D_n^N = \bigcap_{m=n}^N E_m^c$ for $N \geq n \geq 1$.

(c) Show that $\mathbb{P}(D_n^N) \leq \exp(-\sum_{m=n}^N \mathbb{P}(E_m))$. [*Hint: it holds that* $1 - x \leq e^{-x}$.]

(d) Let $D_n^\infty := \bigcap_{m=n}^\infty E_m^c$. Show that $\mathbb{P}(D_n^\infty) = 0$.

(e) Show that $\mathbb{P}(\liminf E_n^c) = 0$.

(f) Show that $\mathbb{P}(\limsup E_n) = 1$.

25. Let $X_1, X_2, \ldots$ be a sequence of *nonnegative* random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and put $S_\infty = \sum_{i=1}^\infty X_i$. We also have the measurable space $(\mathbb{N}, \mathcal{N}, \tau)$, where $\mathbb{N}$ is the set of positive integers, $\mathcal{N}$ is the power set of $\mathbb{N}$ and $\tau$ the counting measure.

We consider the product set $\mathbb{N} \times \Omega$ with the product $\sigma$-algebra $\mathcal{N} \times \mathcal{F}$ and the product measure $\tau \times \mathbb{P}$. On the product set we define the mapping $\mathbf{X} : \mathbb{N} \times \Omega \to \mathbb{R}$ by $\mathbf{X}(k, \omega) = X_k(\omega)$. Let, for a given Borel set $B$ in $\mathbb{R}$, $A := \mathbf{X}^{-1}[B] = \{(k, \omega) : \mathbf{X}(k, \omega) \in B\}$ and $A_k := X_k^{-1}[B] = \{\omega : X_k(\omega) \in B\}$ , for $k \in \mathbb{N}$. Note that $A = \cup_{k \in \mathbb{N}}(\{k\} \times A_k)$, i.e. $(k, \omega) \in A$ iff $\omega \in A_k$.

(a) Why are the sets $\{k\} \times A_k$ above elements of $\mathcal{N} \times \mathcal{F}$?

(b) Show that $\mathbf{X}$ is a measurable mapping on $\mathbb{N} \times \Omega$ with the product $\sigma$-algebra $\mathcal{N} \times \mathcal{F}$, i.e. the set $A$ above belongs to $\mathcal{N} \times \mathcal{F}$ (for any Borel set $B$).

(c) Show by an application of Fubini's theorem (recall that summation is an example of Lebesgue integration) that $\mathbb{E}\, S_\infty = \sum_{i=1}^\infty \mathbb{E}\, X_i$.

(d) If the $X_i$ are not necessarily nonnegative, give then an integrability condition on the $X_i$ such that the equality $\mathbb{E}\, S_\infty = \sum_{i=1}^\infty \mathbb{E}\, X_i$ is still true.

26. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined *nonnegative* random variables $X$, $X_n$ $(n \geq 1)$ that have the property that $X_n \xrightarrow{\mathbb{P}} X$ (so $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$ as $n \to \infty$ for every $\varepsilon > 0$). Let $Y_n = \frac{X_n}{1+X_n}$ and $Y = \frac{X}{1+X}$ and note that $Y_n \leq 1$.

(a) Do we have $Y_n \xrightarrow{\mathbb{P}} Y$?

(b) Show the two inequalities $|Y_n - Y| \leq 2$ and $|Y_n - Y| \leq |X_n - X|$.

(c) Show that $\mathbb{E}\, |Y_n - Y| \leq 2\, \mathbb{P}(|X_n - X| > \varepsilon) + \varepsilon$ for every $\varepsilon > 0$.

(d) Show that $Y_n \xrightarrow{\mathcal{L}^1} Y$, i.e. $\mathbb{E}\, |Y_n - Y| \to 0$.

27. Consider a sample from exponential distribution, i.e. one has an IID sequence $X_1, \ldots, X_n$ where all $X_i$ have a density $p_\lambda(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ and a parameter $\lambda > 0$. Along with the $X_i$ one also observes $Y_i = \cos X_i$, $i = 1, \ldots, n$. Probabilities, expectations, etc. depending on $\lambda$, when necessary, are denoted $\mathbb{P}_\lambda$, $\mathbb{E}_\lambda$, etc. and $\overline{Y}_n$ is the average of the $Y_i$.

(a) Show that $\mathbb{E}_\lambda Y_i = \frac{\lambda^2}{1+\lambda^2}$ [*Hint: use two times integration by parts, for which you may want to use that* $\frac{\mathrm{d} \sin x}{\mathrm{d}x} = \cos x$ *and* $\frac{\mathrm{d} \cos x}{\mathrm{d}x} = -\sin x$.]

(b) Show that the moment estimator using the $Y_i$ as (transformed) observations is $\hat{\lambda}_n = \sqrt{\frac{\overline{Y}_n}{1-\overline{Y}_n}}$, provided that $\overline{Y}_n \in [0, 1)$.

(c) Show that $\overline{Y}_n < 1$ a.s. and show by invoking the Law of Large Numbers (LLN) for $\overline{Y}_n$ that $\overline{Y}_n > 0$ with probability tending to 1.

(d) Show by using the above LLN that $\hat{\lambda}_n \xrightarrow{\mathbb{P}_\lambda} \lambda$ (so, the $\hat{\lambda}_n$ are consistent estimators of $\lambda$).

The standardized moment estimator $\sqrt{n}(\hat{\lambda}_n - \lambda)$ has a limit law, which is normal with variance $\frac{(1+\lambda^2)^4}{4\lambda^2}\sigma^2(\lambda)$, where $\sigma^2(\lambda)$ is $\mathbb{V}\mathrm{ar}_\lambda(Y_1)$. [*In fact* $\sigma^2(\lambda) = \frac{5\lambda^2+2}{(\lambda^2+4)(\lambda^2+1)^2}$, *which we take for granted.*] Below you are asked to provide two justifications of this result.

(e) Show by application of the theory of moment estimators that the postulated limit law is correct. [*If it is convenient for you, you rename the above $\lambda$ as $\lambda_0$, the 'true' parameter.*]

(f) Show by application of the theory for M-estimators that the postulated limit law is correct. [*You don't have to verify the conditions of the theorem you'd like to use; just blindly apply the assertions.*]

28. Let $\Omega = [0,1]$ and define a collection of subsets of $\Omega$, call it $\mathcal{F}$, by $F \in \mathcal{F}$ if either $F$ is at most countable or its complement $F^c$ is at most countable. Furthermore, let the map $\mathbb{P} : \mathcal{F} \to [0,1]$ be defined by

$$\mathbb{P}(F) = \begin{cases} 0 \text{ if } F \text{ is at most countable} \\ 1 \text{ if } F^c \text{ is at most countable,} \end{cases}$$

and note that we don't a priori impose that $\mathbb{P}$ is a measure.

(a) Show that $\mathcal{F}$ is a $\sigma$-algebra.

(b) Let $A_1, A_2, \ldots$ be a disjoint sequence in $\mathcal{F}$. Show that $\mathbb{P}(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mathbb{P}(A_i)$ if all $A_i$ are at most countable, or if there is exactly one $A_i$ whose complement is at most countable.

(c) Suppose that $A_1, A_2 \in \mathcal{F}$ such that $A_1^c$ and $A_2^c$ are at most countable. Are $A_1$ and $A_2$ disjoint?

(d) Is $\mathbb{P}(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mathbb{P}(A_i)$ for all disjoint sequences $A_1, A_2, \ldots$ in $\mathcal{F}$?

(e) Is $\mathbb{P}$ a probability measure on $\mathcal{F}$?

29. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X_1, X_2, \ldots$ be an IID sequence of random variables defined on it with $\mathbb{E}\, X_1^2 < \infty$, $\mathbb{E}\, X_1 = \mu$ and $\operatorname{Var} X_1 = \sigma^2$. Furthermore $N : \Omega \to \{1, 2, \ldots\}$ is a random variable on this space that is independent of the $X_i$, with $\mathbb{E}\, N^2 < \infty$, $\mathbb{E}\, N = m$ and $\operatorname{Var} N = v^2$. Put $S_n = \sum_{i=1}^n X_i$, $S = \sum_{i=1}^N X_i$ and note that the number of terms in the latter sum is random. Observe that $S = \sum_{n=1}^\infty \mathbf{1}_{\{N=n\}} S_n = \sum_{n=1}^\infty \sum_{i=1}^n \mathbf{1}_{\{N=n\}} X_i$, $|S| \le \sum_{n=1}^\infty \sum_{i=1}^n \mathbf{1}_{\{N=n\}} |X_i|$. Moreover $S^2 \le N \sum_{i=1}^N X_i^2$.

(a) Show that $\mathbb{E}\,|S| < \infty$ and $\mathbb{E}\, S^2 \le \mathbb{E}\, N^2 \times \mathbb{E}\, X_1^2 < \infty$.

(b) Show that $\mathbb{E}\, S = m\mu$.

(c) Show that $\operatorname{Var} S = m\sigma^2 + \mu^2 v^2$.

30. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with a sequence $(\mathcal{F}_n)_{n=1}^{\infty}$ of sub-$\sigma$-algebras of $\mathcal{F}$ satisfying $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for all $n \geq 0$. Such a sequence is called a *filtration*. Furthermore, there is sequence of *independent* random variables $(X_i)_{i=1}^{\infty}$ such that $\mathbb{E}\,|X_i| < \infty$ for all $i$, and every $X_i$ is $\mathcal{F}_i$-measurable. Put $S_n = \sum_{i=1}^{n} X_i$ for $n \geq 1$, $S_0 = 0$, and $P_n = \prod_{i=1}^{n} X_i$ for $n \geq 1$, $P_0 = 1$.

    (a) Show that $S_n$ and $P_n$ are $\mathcal{F}_n$-measurable for all $n$.

    (b) Show that $\mathbb{E}\,|S_n| < \infty$ for all $n$.

    (c) Show that $\mathbb{E}\,[S_{n+1}|\mathcal{F}_n] = S_n$ for all $n \geq 0$ if all $X_n$ have zero expectation.

    (d) Show that $\mathbb{E}\,|P_n| < \infty$ for all $n$. [*You may want to use induction here.*]

    (e) Show that $\mathbb{E}\,[P_{n+1}|\mathcal{F}_n] = P_n$ for all $n \geq 0$ if all $X_n$ have expectation equal to 1.

    NB: The sequences $(S_n)$ are $(P_n)$ are known as examples of *martingales*.

31. Let $X, X_1, X_2, \ldots$ be random variables and suppose $X_n \rightsquigarrow X$. Let $F$ denote the distribution function of $X$.

    (a) Show that $\limsup_{n\to\infty} \mathbb{P}(X_n < x) \leq F(x)$ for all $x$ at which $F$ is continuous.

    (b) Show that $\liminf_{n\to\infty} \mathbb{P}(X_n < x) \geq F(x)$ for all $x$ at which $F$ is continuous. [*Here* $\mathbb{P}(X_n < x) \geq \mathbb{P}(X_n \leq x - \varepsilon)$ *for any* $\varepsilon > 0$ *comes in handy.*

    (c) Conclude that $\lim_{n\to\infty} \mathbb{P}(X_n < x) = F(x)$ for all $x$ at which $F$ is continuous.

32. Consider a sample from an exponential distribution, i.e. one has an IID sequence $X_1, \ldots, X_n$ where all $X_i$ have a density $p_\lambda(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ and a parameter $\lambda > 0$. Along with the $X_i$ one also observes $Y_i = \sin X_i$, $i = 1, \ldots, n$. Probabilities, expectations, etc. depending on $\lambda$, when necessary, are denoted $\mathbb{P}_\lambda$, $\mathbb{E}_\lambda$, etc., and $\overline{Y}_n$ is the average of the $Y_i$.

    (a) Show that $h(\lambda) := \mathbb{E}_\lambda Y_i = \frac{\lambda}{1+\lambda^2}$ [*Hint: use two times integration by parts, for which you may want to use that* $\frac{\mathrm{d}\sin x}{\mathrm{d}x} = \cos x$ *and* $\frac{\mathrm{d}\cos x}{\mathrm{d}x} = -\sin x$.]

    (b) Show that $h(\lambda) \leq \frac{1}{2}$ with equality iff $\lambda = 1$, and $h(\frac{1}{\lambda}) = h(\lambda)$. [Note that therefore $\lambda$ cannot be identified from $\mathbb{E}_\lambda Y_i$.]

(c) Show, use the law of large numbers, that $\mathbb{P}_\lambda(\overline{Y}_n \leq \frac{1}{2}) \to 1$ for $\lambda \neq 1$.

(d) Show that a possible moment estimator using the $Y_i$ as (transformed) observations is $\hat{\lambda}_n = \frac{1}{2\overline{Y}_n}(1 \pm \sqrt{1 - 4(\overline{Y}_n)^2}) =: g_\pm(\overline{Y}_n)$.

(e) Choose $\hat{\lambda}_n = \frac{1}{2\overline{Y}_n}(1 + \sqrt{1 - 4(\overline{Y}_n)^2}) =: g_+(\overline{Y}_n)$ as the estimator of $\lambda$. Show that $\hat{\lambda}_n \xrightarrow{\mathbb{P}_\lambda} g(\lambda)$, where $g(\lambda) = \lambda \mathbf{1}_{\lambda > 1} + \frac{1}{\lambda}\mathbf{1}_{\lambda \leq 1} = \max\{\lambda, \frac{1}{\lambda}\}$. [The $\hat{\lambda}_n$ are consistent estimators of $g(\lambda)$.]

To avoid identification and other technical problems, we assume that it is known that $\lambda > 1$. As a result $\hat{\lambda}_n \xrightarrow{\mathbb{P}_\lambda} \lambda$. The standardized moment estimator $\sqrt{n}(\hat{\lambda}_n - \lambda)$ has a limit law, which is normal with variance $\frac{(1+\lambda^2)^4}{(1-\lambda^2)^2}\sigma^2(\lambda)$, where $\sigma^2(\lambda)$ is $\mathbb{V}\mathrm{ar}_\lambda(Y_1)$, which we don't compute. Below you are asked to provide two justifications of this result.

(f) Show by application of the theory of moment estimators, and check that the relevant conditions are satisfied, that the postulated limit law is correct. [*If it is convenient for you, you can rename the above $\lambda$ as $\lambda_0$, the 'true' parameter.*]

(g) Show by application of the theory for M-estimators that the postulated limit law is correct. [*You don't have to verify the conditions of the theorem you'd like to use; just blindly apply the assertions.*]

33. We consider a sample from an inverse Gamma distribution. We have nonnegative IID observations $X_i$ with a common density, given by $p_\beta(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp(-\beta/x)$ where $\alpha > 0$ is known and $\beta > 0$ a parameter to be estimated. In order to do that we consider a moment estimator with the function $f(x) = \frac{1}{x}$, that is we solve the equation $\mathbb{E}_\beta f(X_1) = \frac{1}{n}\sum_{i=1}^n f(X_i)$, in alternative notation $P_\beta f = \mathbb{P}_n f$. Some additional information is $\mathbb{V}\mathrm{ar}_\beta \frac{1}{X_1} = \frac{\alpha}{\beta^2}$, and $\mathbb{V}\mathrm{ar}_\beta X_1 = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ if $\alpha > 2$.

(a) Show that $\mathbb{E}_\beta \frac{1}{X_1} = \frac{\alpha}{\beta}$, and $\mathbb{E}_\beta X_1 = \frac{\beta}{\alpha - 1}$ if $\alpha > 1$. [*Recall that* $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.]

(b) Show that the moment estimator of $\beta$ with the given $f$ as above is $\hat{\beta}_n = \frac{n\alpha}{\sum_{i=1}^n \frac{1}{X_i}}$.

(c) Find the limit law of $\sqrt{n}(\hat{\beta}_n - \beta)$. [*You may want to use the ordinary Central Limit Theorem as a first step.*]

(d) An alternative moment estimator, call it $\hat{\beta}'_n$, is obtained by solving the equation $\frac{1}{n}\sum_{i=1}^n X_i = \mathbb{E}_\beta X_1$. What is the asymptotic distribution of this estimator?

(e) Which of the two estimators $\hat{\beta}_n$, $\hat{\beta}'_n$ is preferred, both from a 'quality' point of view –think of some desirable properties as consistency, efficiency, (asymptotic) mean squared error– and from the point of view of weaker or stronger underlying assumptions?

(f) Is there an estimator that performs (strictly) better than the most preferred one in (e) from the 'quality' point of view?

34. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which random variables $X$, $Y$ are defined. It is assumed that the vector $(X, Y)$ has a joint density w.r.t. Lebesgue measure, a nonnegative Borel-measurable function $f : \mathbb{R}^2 \to \mathbb{R}$, integrating to 1. Moreover, $f$ is such that the marginal density $f_Y$ of $Y$ is strictly positive. One can now define the conditional density of $X$ given $Y = y$, $f_{X|Y=y}(x) := \frac{f(x,y)}{f_Y(y)}$ for all $x, y$, where $f_Y$ is the marginal density of $Y$. It is assumed that $\mathbb{E}\,|X| < \infty$. Integrals with the densities can also be viewed as Riemann integrals.

    (a) Show that $f_{X|Y=y}$ is a density for all $y$.

    (b) Let $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(y) := \int_{\mathbb{R}} x f_{X|Y=y}(x)\,\mathrm{d}x$. The theory for product measures says that $g$ is a Borel function. Show that $g(Y)$ is a proper random variable, i.e. a function on $\Omega$ that is $\mathcal{F}$-measurable.

    (c) Let $G = \{Y \in B\}$, where $B$ is a Borel set in $\mathbb{R}$. Show by representing the expectations as (double) integrals that $\mathbb{E}\left[\mathbf{1}_G g(Y)\right] = \mathbb{E}\left[\mathbf{1}_G X\right]$.

    (d) What is $\mathbb{E}\left[X|Y\right]$ in terms of $g$?

35. Let $f$, $f_n$ $(n \geq 1)$ be densities of probability measures on $(\mathbb{R}, \mathcal{B})$, nonnegative Borel-measurable function such that their integrals w.r.t Lebesgue measure (denoted $\lambda$) equal 1. By $F$, $F_n$ we denote the corresponding distribution functions. It is assumed that $f_n \to f$ a.e. (w.r.t Lebesgue measure), as $n \to \infty$.

    (a) Show that $F$ is everywhere *left* continuous, i.e. $\lim_{n\to\infty} F(x - \frac{1}{n}) = F(x)$ for all $x \in \mathbb{R}$. [*Consider* $\int \mathbf{1}_{(x-\frac{1}{n},x]} f \,\mathrm{d}\lambda$.]

    (b) Show that $\int |f - f_n|\,\mathrm{d}\lambda = 2\int \mathbf{1}_{\{f>f_n\}}(f - f_n)\,\mathrm{d}\lambda$

    (c) Show that $0 \leq \mathbf{1}_{\{f>f_n\}}(f - f_n) \leq f$ and deduce that $\int |f - f_n|\,\mathrm{d}\lambda \to 0$.

(d) Show that $F_n \rightsquigarrow F$, i.e. $F_n(x) \to F(x)$ for all $x \in \mathbb{R}$. [*Use the previous question and work on* $|F_n(x) - F(x)|$.] Why do we have convergence for all $x$?

36. Let $X$ have a Binomial distribution with parameters $n$ and $p$, $p$ is unknown. Independent of $X$, there is a random variable $Y$ having a Binomial distribution with parameters $m$ and $q$, where $q > 0$ is also unknown. Assume $m = 2n$. We are interested to estimate $r := \frac{p}{q}$ and in asymptotic behavior of estimators for $n \to \infty$.

    (a) Give a consistent estimator $\hat{r}_n$ of $r$, and show its consistency.

    (b) Give the limit distribution of $\sqrt{n}(\hat{r}_n - r)$.

    (c) It may happen that although $p$ and $q$ are unknown, it is known that they are equal. Give a better estimator of $r$ than the $\hat{r}_n$ above.

37. Let $X_1, \ldots, X_n$ be independent random variables with a common $N(\theta_0, 1)$ distribution with $\theta_0 \in \mathbb{R}$. As an estimator $\hat{\theta}_n$ we use the minimizer of the function $\theta \mapsto \sum_{i=1}^n (X_i - \theta)^6$. This minimizer also solves an equation of the type $\Psi_n(\theta) = 0$.

    (a) Show that a well chosen $\Psi_n(\theta)$ converges in probability to a limit $\Psi(\theta)$ and determine this limit. Show, compute a derivative, that $\Psi(\theta)$, is a strictly monotone function of $\theta$.

    (b) Show that $\hat{\theta}_n$ is a consistent estimator of $\theta_0$.

    (c) Show that $\sqrt{n}(\hat{\theta}_n - \theta_0)$ has a limiting normal distribution and determine its expectation and variance. [Skip the verification of all conditions of a theorem you want to use.]

    *You may want to use that for a standard normal random variable $Z$ it holds that $\mathbb{E}\, Z^p = 0$ for an odd integer $p$ and for an even integer $p = 2m$ one has $\mathbb{E}\, Z^{2m} = 1 \cdot 3 \cdots (2m-1)$, a product of odd integers.*

38. A density of a 1-dimensional exponential family is of the form

$$p_\theta(x) = c(\theta) h(x) e^{\theta t(x)},$$

where $h$ and $t$ are known measurable functions on $\mathbb{R}$, and $h$ is non-negative. It is also assumed that $c$ is a known function of $\theta \in \Theta$. Many well known distributions are of this type. Let $\Theta = \{\theta \in \mathbb{R} : \int_{\mathbb{R}} h(x) e^{\theta t(x)}\, \mathrm{d}x < \infty\}$. Then for $\theta \in \Theta$ one has

$$c(\theta) = \frac{1}{\int_{\mathbb{R}} h(x) e^{\theta t(x)}\, \mathrm{d}x}. \tag{1}$$

Moreover, we can define, for $\theta \in \Theta$,

$$b(\theta) = \log \int_{\mathbb{R}} h(x) e^{\theta t(x)} \, \mathrm{d}x,$$

as the integral is positive. In what follows we denote differentiation w.r.t. $\theta$ by a dot. Integration and differentiation can be swapped at will.

(a) Why does Equation (**??**) hold?

(b) Show that $\Theta$ is an interval, i.e. if $\theta_1, \theta_2 \in \Theta$, then also $\frac{1}{2}(\theta_1 + \theta_2) \in \Theta$.

(c) Show that $\dot{b}(\theta) = \mathbb{E}_\theta t(X)$ (assuming that both expressions exist as finite quantities).

(d) Show that $\ddot{b}(\theta) = \mathbb{V}\mathrm{ar}_\theta t(X)$ (assuming that both expressions exist as finite quantities).

Let now $X_1, \ldots, X_n$ be an IID sample from such an exponential family.

(e) Use $\bar{t}_n := \frac{1}{n} \sum_{i=1}^n t(X_i)$ as a statistic to derive a moment estimator of $\theta$ and write explicitly the equation (in terms of the function $c$ and its derivative) $\theta$ has to solve to get this moment estimator, call it $\hat{\theta}_n^{\mathrm{mom}}$.

(f) Find, informally, the asymptotic distribution of $\hat{\theta}_n^{\mathrm{mom}}$ for $n \to \infty$.

(g) Also write down the equation that the Maximum likelihood estimator, call it $\hat{\theta}_n^{\mathrm{ml}}$, has to satisfy.

(h) Compute the Fisher information $I_\theta$ in one observation.

(i) Find, informally, the asymptotic distribution of $\hat{\theta}_n^{\mathrm{ml}}$ for $n \to \infty$.