# Princeton University

# Spring 2025 MAT425: Measure Theory

# HW9 Sample Solutions

# Apr 28 2025

Serban Cicortas (only Q7) and Chayim Lowen (all other Questions)

May 5, 2025

## Problem 1.

Show that convergence in total variation implies convergence in distribution for random variables.

*Solution.* Let $(\Omega, \mathfrak{M}, \mathbb{P})$ be a probability space. Suppose $X_n : \Omega \to \mathbb{C}$ are random variables for $n \in \mathbb{N}$ and that $X_n \longrightarrow X$ in total variation. This means that the total variation of $\mathbb{P}_{X_n} - \mathbb{P}_X$ goes to 0 as $n \longrightarrow \infty$.[1]

Let $f : \mathbb{C} \to \mathbb{C}$ be a bounded continuous function. We show that $\mathbb{E}[f(X_n)] \longrightarrow E[f(X)]$ as $n \longrightarrow \infty$:

$$\begin{aligned}
|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\equiv \left| \int_\Omega f \circ X_n \, d\mathbb{P} - \int_\Omega f \circ X \, d\mathbb{P} \right| \\
&= \left| \int_\mathbb{C} f \, d\mathbb{P}_{X_n} - \int_\mathbb{C} f \, d\mathbb{P}_X \right| \\
&= \left| \int_\mathbb{C} f \, d(\mathbb{P}_{X_n} - \mathbb{P}_X) \right| \\
&\le (\sup_\mathbb{C} |f|) \cdot \int_\mathbb{C} 1 \, d\, |\mathbb{P}_{X_n} - \mathbb{P}_X| \\
&= (\sup_\mathbb{C} |f|) \cdot \|\mathbb{P}_{X_n} - \mathbb{P}_X\|_{TV}
\end{aligned}$$

Since the last term goes to 0 with $n$, so does the first term. □

[1] Convergence here is with respect to the the total variation norm, namely $\|\cdot\|_{TV} : \mu \mapsto |\mu|(\Omega)$. A related norm is given by $\|\cdot\|_1 := \mu \mapsto \sup_{S \in \mathfrak{M}} |\mu(S)|$. The two norms are equivalent in the sense that $\|\mu\|_1 \le \|\mu\|_{TV} \le \pi \|\mu\|_1$. In particular, the notions of convergence for these two norms coincide.

## Problem 2.

Show that convergence in probability implies convergence in distribution for random variables.

*Solution.* Let $(\Omega, \mathfrak{M}, \mathbb{P})$ be a probability space. Suppose $X_n : \Omega \to \mathbb{C}$ are random variables for $n \in \mathbb{N}$ and that $X_n \longrightarrow X$ in probability. This means we have $\lim_{n\to\infty} \mathbb{P}[|X_n - X| > \epsilon] \longrightarrow 0$ as $n \longrightarrow 0$ for all $\epsilon > 0$.

Let $f : \mathbb{C} \to \mathbb{C}$ be a bounded continuous function. We need to show that $\mathbb{E}[f(X_n)] \longrightarrow E[f(X)]$ as $n \longrightarrow \infty$. It suffices to show that $\limsup_{n\to\infty} |\mathbb{E}[f(X_n) - f(X)]| \leq \epsilon$ for arbitrary $\epsilon > 0$. Fixing $\epsilon > 0$, we define for each $n$

$$B_n := \{\omega \in \Omega \mid |f(X_n(\omega)) - f(X(\omega))| > \epsilon\}$$

We claim that $\mathbb{P}(B_n) \longrightarrow 0$ as $n \longrightarrow \infty$. Assuming this, it follows that

$$\begin{aligned}
|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &= |\mathbb{E}[f \circ X_n - f \circ X]| \\
&= \left| \int_\Omega (f \circ X_n - f \circ X) \, d\mathbb{P} \right| \\
&\leq \int_\Omega |f \circ X_n - f \circ X| \, d\mathbb{P} \\
&= \int_{B_n} |f \circ X_n - f \circ X| \, d\mathbb{P} + \int_{\Omega - B_n} |f \circ X_n - f \circ X| \\
&\leq \int_{B_n} (|f \circ X_n| + |f \circ X|) + \int_{\Omega - B_n} \epsilon \\
&\leq (2 \sup_{\mathbb{C}} f) \cdot \mathbb{P}(B_n) + \epsilon
\end{aligned}$$

Taking lim sup of the first and last terms of this inequality gives the desired result.

It remains to show that $\mathbb{P}[B_n] \longrightarrow 0$ as $n \longrightarrow \infty$ (i.e. that $f(X_n) \longrightarrow f(X)$ in probability). To see this, fix $\epsilon > 0$ and choose $M > 0$ such that $\mathbb{P}[|X| < M] > 1 - \epsilon$. The uniform continuity of $f$ on the closed disk of radius $M + 1$ implies that for some $\delta \in (0, 1)$, we have $|f(x) - f(y)| < \epsilon$ for all $x, y \in \mathbb{C}$ satisfying $|x|, |y| < M + 1$ and $|x - y| < \delta$. Without loss of generality, we may suppose that $\mathbb{P}[|X_n - X| < \delta] < \epsilon$ for all $n$.[2] Then $|f(X_n(\omega)) - f(X(\omega))| \geq \epsilon$ implies that either $|X_n(\omega) - X(\omega)| > \delta$ or $|X_n(\omega)| > M + 1$ or $|X(\omega)| > M + 1$. If $|X_n(\omega) - X(\omega)| < \delta$, then the last two cases are covered by the case that $|X(\omega)| > M$. Putting all this together gives

$$\mathbb{P}[B_n] = \mathbb{P}[|f(X_n) - f(X)| > \epsilon] \leq \mathbb{P}[|X_n - X| > \delta] + \mathbb{P}[|X| > M] \leq 2\epsilon$$

The result now follows from the fact that $\epsilon$ can be made arbitrarily small. □

[2] This can be achieved by restricting to a tail of the sequence $f_1, f_2, \ldots$.

## Problem 3.

Let $(\Omega, \mathfrak{M}, \mathbb{P})$ be a probability space. Suppose $X_n : \Omega \to \mathbb{R}$ for $n \in \mathbb{N}$ and and $X : \Omega \to \mathbb{R}$ are (real-valued) random variables. Show that $X_n \longrightarrow X$ in distribution if and only if $\mathbb{P}[X_n < t] \longrightarrow \mathbb{P}[X < t]$ for all values of $t \in \mathbb{R}$ at which $x \mapsto \mathbb{P}[X < x]$ is continuous. C.f. Theorem 2.3 in Varadhan's *Probability Theory.*

*Solution.* $\Longrightarrow$ Fix $t \in \mathbb{R}$ at which $x \mapsto \mathbb{P}[X < x]$ is continuous. Let $\epsilon > 0$. Choose $\delta > 0$ such that

$$\mathbb{P}[X < t - \delta] > \mathbb{P}[X > t] - \epsilon, \qquad \mathbb{P}[X < t + \delta] < \mathbb{P}[X < t] + \epsilon$$

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function satisfying[3]

(i) $f(x) = 1$ for $x \leq t - \delta$.

(ii) $f(x) = 0$ for $x \geq t$.

(iii) $0 \leq f(x) \leq 1$ for all $x \in \mathbb{R}$.

Then

$$\mathbb{P}[X_n < t] \geq \mathbb{E}[f(X_n)] \longrightarrow \mathbb{E}[f(X)] \geq \mathbb{P}[X < t - \delta] > \mathbb{P}[X < t] - \epsilon$$

It follows that

$$\liminf_{n\to\infty} \mathbb{P}[X_n < t] \geq \mathbb{P}[X < t] - \epsilon$$

Similarly, let $g : \mathbb{R} \to \mathbb{R}$ be a continuous function satisfying

(i) $g(x) = 1$ for $x \leq t$.

(ii) $g(x) = 0$ for $x \geq t + \delta$.

(iii) $0 \leq g(x) \leq 1$ for all $x \in \mathbb{R}$.

Then

$$\mathbb{P}[X_n < t] \leq \mathbb{E}[g(X_n)] \longrightarrow \mathbb{E}[g(X)] \leq \mathbb{P}[X < t + \delta] < \mathbb{P}[X < t] + \epsilon$$

It follows that

$$\limsup_{n\to\infty} \mathbb{P}[X_n < t] \leq \mathbb{P}[X < t] + \epsilon$$

Since $\epsilon$ can be made arbitrarily small in the inequalities

$$\limsup_{n\to\infty} \mathbb{P}[X_n < t] - \epsilon \leq \mathbb{P}[X < t] \leq \liminf_{n\to\infty} \mathbb{P}[X_n < t] + \epsilon$$

it follows that $\lim_{n\to\infty} \mathbb{P}[X_n < t] = \mathbb{P}[X < t]$.

[3] For instance, we can find a piecewise-linear function with these properties.

$\Longleftarrow$ Note that $x \mapsto \mathbb{P}[X < x]$ is an *increasing* function of $x$. This implies that it has at most countably many points of discontinuity. In particular, its points of continuity are dense in $\mathbb{R}$.

Let $f : \mathbb{R} \to \mathbb{C}$ be a bounded continuous function. Let $K > 0$ be an upper bound on $|f|$. Let $\epsilon > 0$. Choose $M > 0$ such that $\mathbb{P}[X < -M] < \epsilon$ and $\mathbb{P}[X < M] > 1 - \epsilon$. We can additionally choose $M$ such that both $M$ and $-M$ are points of continuity of $x \mapsto \mathbb{P}[X < x]$. Without loss of generality, we may also suppose that $\mathbb{P}[X_n < -M] < 2\epsilon$ and $\mathbb{P}[X_n < M] > 1 - 2\epsilon$ for all $n$.[4]

Since $[-M, M]$ is compact, $f$ is uniformly continuous on this interval. So we can find a sequence $x_0, x_1, \ldots, x_N$ (for some $N$) such that

(i) $x_0 = -M$, $x_N = M$

(ii) $x_0 < x_1 < \cdots < x_N$

(iii) Each point $x_j$ is a point of continuity of $x \mapsto \mathbb{P}[X < x]$

(iv) $|f(y_1) - f(y_2)| < \epsilon$ for all $y_1, y_2$ in a common interval $[x_j, x_{j+1}]$.

Define $g : \mathbb{R} \to \mathbb{C}$ by $g = \sum_{j=1}^{n} f(x_{j-1})\chi_{[x_{j-1}, x_j)}$. The hypotheses above imply that

$$|f(t) - g(t)| \leq \epsilon \quad \text{for } t \in [-M, M] \qquad \text{and} \qquad |f(t) - g(t)| \leq 2K \quad \text{for all } t$$

Hence for all $n$

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[g(X_n)]| \leq \epsilon + K\,\mathbb{P}[X_n < -M] + K\,\mathbb{P}[X_n \geq M] \leq \epsilon + 4K\epsilon$$

Similarly,

$$|\mathbb{E}[f(X)] - \mathbb{E}[g(X)]| \leq \epsilon + K\,\mathbb{P}[X < -M] + K\,\mathbb{P}[X > M] \leq \epsilon + 2K\epsilon$$

Hence, by the triangle inequality,

$$\begin{aligned} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\leq |\mathbb{E}[f(X_n)] - \mathbb{E}[g(X_n)]| + |\mathbb{E}[g(X_n)] - \mathbb{E}[g(X)]| + |\mathbb{E}[g(X)] - \mathbb{E}[f(X)]| \\ &\leq (6K + 2)\epsilon + \sum_{j=1}^{n} f(x_{j-1})\,|\mathbb{P}[x_{j-1} \leq X_n < x_j] - \mathbb{P}[x_{j-1} \leq X < x_j]| \end{aligned}$$

Since

$$\mathbb{P}[x_{j-1} \leq X_n < x_j] - \mathbb{P}[x_{j-1} \leq X < x_j] = (\mathbb{P}[X_n < x_j] - \mathbb{P}[X < x_j]) - (\mathbb{P}[X_n < x_{j-1}] - \mathbb{P}[X < x_{j-1}])$$

this term tends to 0 with $n$. Therefore, taking limit suprema in the inequality above gives

$$\limsup_{n \to \infty} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leq (6K + 2)\epsilon$$

Since $\epsilon$ can be made arbitrarily small, the result follows.

$\square$

[4] This is necessarily true for all large enough $n$. So we can achieved the desired result by replacing our sequence of functions by one of its tails.

## Problem 4.

Let $(\Omega, \mathfrak{M}, \mathbb{P})$ be a probability space. Suppose $X_n : \Omega \to \mathbb{R}$ for $n \in \mathbb{N}$ and and $X : \Omega \to \mathbb{R}$ are (real-valued) random variables satisfying $\mathbb{E}[\exp(itX_n)] \longrightarrow \mathbb{E}[\exp(itX)]$ pointwise for $t \in \mathbb{R}$. Then $X_n \longrightarrow X$ in distribution.

*Solution.* We need to show that $\mathbb{E}[f(X_n)] \longrightarrow \mathbb{E}[f(X)]$ for all bounded continuous functions $f : \mathbb{R} \to \mathbb{C}$. The proof of is given under Theorem 7.32 in the lecture notes in the case where $f$ is the Fourier transform of some $L^1$ function $g : \mathbb{R} \to \mathbb{C}$. Now let $\mathcal{F}$ be the image of the Fourier transform operator $L^1(\mathbb{R} \to \mathbb{C}, \lambda) \longrightarrow C_0(\mathbb{R} \to \mathbb{C})$. Note that $\mathcal{F}$ is dense in $C_0(\mathbb{R} \to \mathbb{C})$.[5] Therefore, for any $f \in C_0(\mathbb{R} \to \mathbb{C})$, we can a sequence $f_1, f_2, \ldots$ in $\mathcal{F}$ such that $f_m \longrightarrow f$ *uniformly*. We have (using the triangle inequality)

$$\begin{aligned}
|(\mathbb{E}[f_m(X_n)] - \mathbb{E}[f_m(X)]) - (\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)])| &\leq |\mathbb{E}[f_m(X_n)] - \mathbb{E}[f(X_n)]| + |\mathbb{E}[f_m(X)] - \mathbb{E}[f(X)]| \\
&= |\mathbb{E}[(f_m - f)(X_n)]| + |\mathbb{E}[(f_m - f)(X)]| \\
&\leq 2\|f_m - f\|_\infty
\end{aligned}$$

and therefore (using the triangle inequality again)

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leq |(\mathbb{E}[f_m(X_n)] - \mathbb{E}[f_m(X)])| + 2\|f_m - f\|_\infty$$

We take limit suprema over $n$, noting that the $f_m \in \mathcal{F}$ implies that the middle term tends to 0. This gives

$$\limsup_n |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leq 2\|f_m - f\|_\infty$$

Letting $m \longrightarrow \infty$ shows that $\mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$. This holds for any $f \in C_0(\mathbb{R} \to \mathbb{C})$. To extend the result to $C_b(\mathbb{R} \to \mathbb{C})$, consider any bounded continuous function $f : \mathbb{R} \to \mathbb{C}$. Fix $\epsilon > 0$. Let $M$ be large enough that $\mathbb{P}[-M < X < M] > 1 - \epsilon$. Let $f_1 : \mathbb{R} \to \mathbb{C}$ be a continuous function satisfying[6]

(i) $f_1$ coincides with $f$ on the interval $[-M-1, M+1]$.

(ii) $f_1$ vanishes outside the interval $[-M-2, M+2]$

(iii) $|f_1(x)| \leq |f(x)|$ for all $x \in \mathbb{R}$.

Then

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[f_1(X_n)]| = |\mathbb{E}[(f - f_1)(X_n)]| \leq \mathbb{E}[|f - f_1| \circ X_n] \leq 2\|f\|_\infty \mathbb{P}[|X_n| > M + 1]$$

To bound the last term, let $g : \mathbb{R} \to \mathbb{R}$ be a compactly supported function satisfying

(i) $g = 1$ on the interval $[-M, M]$.

[5] This follows from the combination of Corollary 8.23 and Proposition 8.17 in Folland's *Real Analysis*.

[6] For instance, such function can be obtained by multiplying $f$ by a suitable compactly supported function valued in $[0, 1]$.

(ii) $g$ vanishes outside the interval $[-M-1, M+1]$

(iii) $0 \le g(x) \le 1$ for all $x \in \mathbb{R}$.

Then

$$\mathbb{P}[-M-1 < X_n < M+1] \ge \mathbb{E}[g(X_n)] \longrightarrow \mathbb{E}[g(X)] \ge \mathbb{P}[-M < X < M] > 1-\epsilon$$

It follows that $\mathbb{P}[|X_n| > M+1] < \epsilon$ for all sufficiently large $n$. So going back to the inequality above shows that for all large enough $n$

$$|\mathbb{E}[f(X_n)] - \mathbb{E}[f_1(X_n)]| \le 2\epsilon \|f\|_\infty$$

Since $\epsilon$ can be made arbitrarily small, the result is obtained. □

# Problem 5.

Let $(\Omega, \mathfrak{M}, \mathbb{P})$ be a probability space and $X$ a real-valued random variable on $\Omega$ such that $a \le X \le b$ almost surely. Show that

$$\mathbb{E}[\exp(tX)] \le \exp\left(t\,\mathbb{E}[X] + \frac{t^2(b-a)^2}{8}\right)$$

*Solution.* Consider the centred random variable $Y := X - \mathbb{E}[X]$. Note that $Y$ takes values in $[-\overline{a}, \overline{b}]$ where $\overline{a}, \overline{b} \ge 0$ are given by $\overline{a} := \mathbb{E}[X] - a$, $\overline{b} := b - \mathbb{E}[Y]$. The above inequality becomes

$$\mathbb{E}[\exp(tY)] \le \exp\left(\frac{t^2(\overline{b}+\overline{a})^2}{8}\right)$$

Using the convexity of the function $y \mapsto \exp(ty)$ on the interval $[-\overline{a}, \overline{b}]$, we can write

$$\exp(ty) \le \frac{y+\overline{a}}{\overline{b}+\overline{a}} \exp(t\overline{b}) + \frac{\overline{b}-y}{\overline{b}+\overline{a}} \exp(-t\overline{a})$$

(Here we use the fact that $y = \frac{y+\overline{a}}{\overline{b}+\overline{a}}\overline{b} + \frac{\overline{b}-y}{\overline{b}+\overline{a}}(-\overline{a})$.) Now we replace $y$ with $Y$ and take expectations (noting $\mathbb{E}[Y] = 0$):

$$\mathbb{E}[\exp(tY)] \le \frac{\mathbb{E}[Y]+\overline{a}}{\overline{b}+\overline{a}} \exp(t\overline{b}) + \frac{\overline{b}-\mathbb{E}[Y]}{\overline{b}+\overline{a}} \exp(-t\overline{a}) = \frac{\overline{b}e^{-t\overline{a}} + \overline{a}e^{t\overline{b}}}{\overline{b}+\overline{a}}$$

We claim that the last term is always bounded above by $\exp\left(\frac{t^2(\overline{b}+\overline{a})^2}{8}\right)$. To see this, we define

$$F(t) := \log \frac{\overline{b}e^{-t\overline{a}} + \overline{a}e^{t\overline{b}}}{\overline{b}+\overline{a}}$$

One computes[7]

$$F(0) = 0, \qquad F'(0) = 0, \qquad F''(t) = \frac{ab(a+b)^2}{(be^{-t(a+b)/2} + ae^{t(a+b)/2})^2}$$

[7]It is helpful to know that $(\log f(t))'' = \frac{f''(t)f(t) - f'(t)^2}{f(t)^2}$.

Using the AM-GM inequality, we get the bound

$$be^{-t(a+b)/2} + ae^{t(a+b)/2} \geq 2\sqrt{ab}$$

Hence

$$F''(t) \leq \frac{ab(a+b)^2}{4ab} = \frac{(a+b)^2}{4ab}$$

By integrating twice (or using Taylor expansions), it follows that

$$F(t) \leq \frac{(a+b)^2}{ab} \cdot \frac{t^2}{2}$$

which completes the proof. $\square$

## Problem 6.

Let $(\Omega, \mathfrak{M}, \mathbb{P})$ be a probability space and $X$ a non-negative $L^2$ random variable on $\Omega$ with positive expected value. Show that

$$\mathbb{P}[X \geq \theta \cdot \mathbb{E}[X]] \geq (1-\theta)^2 \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]} \qquad \text{for all } \theta \in [0,1]$$

*Solution.* Define two new random variables $Y, Z : \Omega \to [0, \infty)$ by

$$Y(\omega) := \begin{cases} X(\omega) & \text{if } X(\omega) < \theta\,\mathbb{E}[X] \\ 0 & \text{otherwise} \end{cases}, \qquad Z(\omega) := \begin{cases} X(\omega) & \text{if } X(\omega) > \theta\,\mathbb{E}[X] \\ 0 & \text{otherwise} \end{cases}$$

Observe that $Y + Z = X$ and $YZ = 0$ (pointwise). Starting with the inequality

$$\mathbb{E}[Z^2 \mid Z \neq 0] \geq \mathbb{E}[Z \mid Z \neq 0]^2$$

and multiplying both sides by $\mathbb{P}[Z \neq 0]^2$, we get

$$\mathbb{E}[Z^2]\,\mathbb{P}[Z \neq 0] = \mathbb{E}[Z^2 \mid Z \neq 0]\,\mathbb{P}[Z \neq 0]^2 \geq (\mathbb{E}[Z \mid Z \neq 0]\,\mathbb{P}[Z \neq 0])^2 = \mathbb{E}[Z]^2$$

Thus

$$\mathbb{E}[X^2]\,\mathbb{P}[X \geq \theta\,\mathbb{E}[X]] = \mathbb{E}[Y^2 + Z^2]\,\mathbb{P}[Z \neq 0] \geq \mathbb{E}[Z^2]\,\mathbb{P}[Z \neq 0] \geq \mathbb{E}[Z]^2 = (\mathbb{E}[X] - \mathbb{E}[Y])^2$$

Since $Y \leq \theta\,\mathbb{E}[X]$ pointwise, we have

$$\mathbb{E}[X] - \mathbb{E}[Y] \geq \mathbb{E}[X] - \theta\,\mathbb{E}[X] = (1-\theta)\,\mathbb{E}[X]$$

Putting things together gives

$$\mathbb{E}[X^2]\,\mathbb{P}[X \geq \theta\,\mathbb{E}[X]] \geq (1-\theta)^2\,\mathbb{E}[X]^2$$ $\square$

## Problem 7.

*Solution.* a). We define the function $F(q) = \log \mathbb{E}[Y^q]$. We compute the derivatives of $F$ :

$$F'(q) = \frac{\mathbb{E}[Y^q \log Y]}{\mathbb{E}[Y^q]} = \mathbb{E}_q[\log Y]$$

$$F''(q) = \frac{\mathbb{E}\big[Y^q (\log Y)^2\big]}{\mathbb{E}[Y^q]} - \left(\frac{\mathbb{E}[Y^q \log Y]}{\mathbb{E}[Y^q]}\right)^2 = \mathrm{Var}_q[\log Y]$$

Taking the logarithm and dividing by $1/r$, our goal is to show the following relation:

$$\frac{1}{s}F(s) - \frac{1}{r}F(r) = \int_0^s \frac{1}{r} f_{r,s}(q) F''(q) dq$$

This is simply a calculus exercise. We compute:

$$\int_0^r \frac{1}{r} f_{r,s}(q) F''(q) dq = \frac{s-r}{rs} \int_0^r q F''(q) dq = \frac{s-r}{rs}\big(rF'(r) - F(r)\big) = \frac{s-r}{s}F'(r) - \frac{1}{r}F(r) + \frac{1}{s}F(r)$$

$$\int_r^s \frac{1}{r} f_{r,s}(q) F''(q) dq = \int_0^r \frac{s-q}{s} F''(q) dq = \frac{r-s}{s}F'(r) - \frac{1}{s}F(r) + \frac{1}{s}F(s)$$

b). We apply the result in part a) and use the additional bounds in the problem to get:

$$\mathbb{E}[Y_n^r] = \big(\mathbb{E}[Y_n^s]\big)^{\frac{r}{s}} \cdot \exp\left( - \int_0^s f_{r,s}(q) \mathrm{Var}_q[\log Y_n] dq\right) \leq C_s^{\frac{r}{s}} \exp\left( - c_s n \int_0^s f_{r,s}(q) dq\right)$$

We compute that:

$$\int_0^s f_{r,s}(q) dq = \frac{r(s-r)}{2}$$

We obtain that the optimal constants are:

$$D_r = \sup_{s\in(r,1)} C_s^{\frac{r}{s}},\ d = \inf_{s\in(r,1)} c_s \frac{r(s-r)}{2}$$

□

## Problem 8.

Let $X$ be a nonnegative random variable on a probability space $(\Omega, \mathfrak{M}, \mathbb{P})$. Show that

$$\mathbb{E}[X^s] = s \int_0^\infty t^{s-1}\, \mathbb{P}[X > t]\, d\lambda(t)$$

for all $s > 0$.

*Solution.* Applying HW4Q6, we have

$$\mathbb{E}[X^s] \equiv \int_\Omega X^s\, d\,\mathbb{P} = \int_0^\infty \mathbb{P}[X^s > x]\, d\lambda(x)$$

Applying the change of coordinates $x = t^s$, $d\lambda(x) = st^{s-1}\, d\lambda(t)$, we get

$$\int_0^\infty \mathbb{P}[X^s > x]\, d\lambda(x) = \int_0^\infty \mathbb{P}[X^s > t^s] \cdot st^{s-1}\, d\lambda(t) = s\int_0^\infty t^{s-1} \cdot \mathbb{P}[X > s]\, d\lambda(t)$$ □

## Problem 9.

Let $X$ be a random variable such that there are $0 < \alpha < a$, $\epsilon \in (0,1)$ and $\beta \in (0,\infty)$ for which

$$\mathbb{P}[|X| < \alpha] \leq \beta\sqrt{\mathbb{P}[X \leq -a]\,\mathbb{P}[X \geq a]} + \epsilon$$

Show that

$$\mathbb{E}[X^2] \geq \frac{1-\epsilon}{1+\frac{1}{2}\beta}\alpha^2$$

*Solution.* Note that the AM-GM inequality gives

$$\sqrt{\mathbb{P}[X \leq -a]\,\mathbb{P}[X \geq a]} \leq \frac{1}{2}(\mathbb{P}[X \leq -a] + \mathbb{P}[X \geq a]) \leq \frac{1}{2}(\mathbb{P}[X \leq -\alpha] + \mathbb{P}[X \geq \alpha]) = \frac{1}{2}\,\mathbb{P}[|X| \geq \alpha]$$

So the inequality in the hypothesis becomes

$$\frac{\beta}{2}\,\mathbb{P}[|X| \geq \alpha] + \epsilon \geq \mathbb{P}[|X| < \alpha] = 1 - \mathbb{P}[|X| \geq \alpha]$$

Solving we get

$$\mathbb{P}[|X| \geq \alpha] \geq \frac{1-\epsilon}{1+\frac{1}{2}\beta}$$

We therefore have

$$\begin{aligned}\mathbb{E}[X^2] &\geq \alpha^2\,\mathbb{P}[X^2 \geq \alpha^2] \\ &= \alpha^2\,\mathbb{P}[|X| \geq \alpha] \\ &\geq \frac{1-\epsilon}{1+\frac{1}{2}\beta}\alpha^2\end{aligned}$$

□

## Problem 10

Let $M$ be an $n \times n$ matroid with values in the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ which is positive-definite as a bilinear operator.

1. Calculate the Gaussian normalization factor

$$Z_A := \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Ax\rangle}\, d\lambda(x)$$

*Solution.* Let $\tilde{A} := \frac{1}{2}(A + A^*)$ be the conjugate-symmetrization of $A$. (We have $A = \tilde{A}$ in case $A$ was already Hermitian. Also note that $A^* = A^T$ in the real case.) Then $\langle x, \tilde{A}x\rangle = \langle x, Ax\rangle$ for all $x \in \mathbb{F}^n$.[8] Then $\tilde{A}$ is unitarily (in the real case: orthogonally) diagonalizable. Let $\nu_1, \ldots, \nu_n$ be the eigenvalues

[8] To see this, note that the hypotheses imply $\langle x, Ax\rangle = \overline{\langle x, Ax\rangle} = \langle Ax, x\rangle = \langle x, A^*x\rangle$.

of $\tilde{A}$. Note that these are positive real numbers. Let $D = \mathrm{diag}(\nu_1, \ldots, \nu_n)$ and let $U$ be a unitary (or orthogonal) matrix such that $\tilde{A} = U^* DU$. Then

$$\left\langle x, \tilde{A}x \right\rangle = \left\langle x, U^* \tilde{D} U x \right\rangle = \langle Ux, DUx \rangle$$

On the other hand, $x \mapsto Ux$ is an isometry of $\mathbb{F}^n$. In particular, it preserves the Lebesgue measure. Thus

$$Z_A \equiv \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Ax \rangle} \, d\lambda(x) = \int_{\mathbb{F}^n} e^{-\frac{1}{2}\left\langle x, \tilde{A}x \right\rangle} \, d\lambda(x) = \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle Ux, DUx \rangle} \, d\lambda(x) = \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Dx \rangle} \, d\lambda(x)$$

Using Tonelli's theorem, we can write

$$\int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Dx \rangle} \, d\lambda(x) = \int_{\mathbb{F}^n} e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \, d\lambda(x) = \prod_{j=1}^n \int_{\mathbb{F}} e^{-\frac{1}{2}\nu_j |x_j|^2} \, d\lambda(x_j)$$

Using a (simple) change-of-coordinates formula, we get

$$\int_{\mathbb{F}} e^{-\frac{1}{2}\nu_j |t|^2} \, d\lambda(t) = \nu_j^{-d/2} \int_{\mathbb{F}} e^{-\frac{1}{2}|t|^2} \, d\lambda(t)$$

where $d = 1$ if $\mathbb{F} = \mathbb{R}$ and $d = 2$ if $\mathbb{F} = \mathbb{C}$. The Gaussian integral on the right has value $(2\pi)^{d/2}$. From this we get

$$\int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Dx \rangle} \, d\lambda(x) = \left( \prod_{j=1}^n \frac{2\pi}{\nu_j} \right)^{d/2} = \frac{(2\pi)^{nd/2}}{(\det \tilde{A})^{d/2}}$$

□

2. Evaluate the integral

$$\int_{x \in \mathbb{F}^n} e^{-\frac{1}{2}\langle x, Ax \rangle + \langle v, x \rangle} \, d\lambda(x)$$

*Solution.* We deal only with the case $\mathbb{F} = \mathbb{R}$. Again, we replace $A$ with $\tilde{A}$ with no effect on the integral. We have

$$\begin{aligned}
\int_{x \in \mathbb{R}^n} e^{-\frac{1}{2}\langle x, Ax \rangle + \langle v, x \rangle} \, d\lambda(x) &= e^{\frac{1}{2}\left\langle \tilde{A}^{-1} v, v \right\rangle} \int_{x \in \mathbb{R}^n} e^{-\frac{1}{2}\left\langle (x - \tilde{A}^{-1} v, \, \tilde{A}(x - \tilde{A}^{-1} v) \right\rangle} \, d\lambda(x) \\
&= e^{\frac{1}{2}\left\langle \tilde{A}^{-1} v, v \right\rangle} \int_{x \in \mathbb{R}^n} e^{-\frac{1}{2}\left\langle x, \, \tilde{A}x \right\rangle} \, d\lambda(x) \\
&= e^{\frac{1}{2}\left\langle \tilde{A}^{-1} v, v \right\rangle} \sqrt{\prod_{j=1}^n \frac{2\pi}{\nu_j}}
\end{aligned}$$

In particular, if $A$ is orthogonal then

$$\int_{x \in \mathbb{R}^n} e^{-\frac{1}{2}\langle x, Ax \rangle + \langle v, x \rangle} \, d\lambda(x) = e^{\frac{1}{2}\left\langle A^{-1} v, v \right\rangle} \sqrt{\prod_{j=1}^n \frac{2\pi}{\nu_j}}$$

□

3. For $v_1, v_2 \in \mathbb{F}^n$, evaluate the expectation value

$$Z_A^{-1} \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Ax\rangle} \langle v_1, x\rangle \langle x, v_2\rangle \; d\lambda(x)$$

*Solution.* We again replace $A$ with $\tilde{A}$. We also write $\tilde{A} = U^*DU$ for unitary $U$ and diagonal $D$.

$$\begin{aligned}
\int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, \tilde{A}x\rangle} \langle v_1, x\rangle \langle x, v_2\rangle \; d\lambda(x) &= \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle Ux, DUx\rangle} \langle v_1, x\rangle \langle x, v_2\rangle \; d\lambda(x) \\
&= \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Dx\rangle} \langle v_1, U^*x\rangle \langle U^*x, v_2\rangle \; d\lambda(x) \\
&= \int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Dx\rangle} \langle Uv_1, x\rangle \langle x, Uv_2\rangle \; d\lambda(x)
\end{aligned}$$

Then writing $Uv_1 = (u_1, \ldots, u_n)$ and $Uv_2 = (w_1, \ldots, w_n)$, we have

$$\begin{aligned}
\int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, Dx\rangle} \langle Uv_1, x\rangle \langle x, Uv_2\rangle \; d\lambda(x) &= \int_{\mathbb{F}^n} e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \sum_{1\le j,k\le n} \overline{u_j} w_k x_j \overline{x_k} \, d\lambda(x) \\
&= \sum_{1\le j,k\le n} \int_{\mathbb{F}^n} e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \overline{u_j} w_k x_j \overline{x_k} \, d\lambda(x)
\end{aligned}$$

We decompose this last term into two kinds of summands:

$$\sum_{j=1}^n \int_{\mathbb{F}^n} e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \overline{u_j} w_j \left|x_j\right|^2 \; d\lambda(x) + \sum_{\substack{1\le j,k\le n \\ j\neq k}} \int_{\mathbb{F}^n} e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \overline{u_j} w_k x_j \overline{x_k} \, d\lambda(x)$$

Each term in the second sum vanishes because each of the integrands $e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \overline{u_j} w_k x_j \overline{x_k}$ is *odd* in the variable $x_j$. Each term in the first sum can be rewritten (using Fubini-Tonelli) as

$$\begin{aligned}
\int_{\mathbb{F}^n} e^{-\frac{1}{2}\sum_{j=1}^n \nu_j |x_j|^2} \overline{u_j} w_j \left|x_j\right|^2 \; d\lambda(x_j) &= \left( \prod_{\substack{1\le k\le n \\ k\neq j}} \int_{\mathbb{F}} e^{-\nu_k |x_k|^2/2} \, d\lambda(x_k) \right) \cdot \overline{u_j} w_j \int_{\mathbb{F}} e^{-\nu_j |x_j|^2/2} \left|x_j\right|^2 \; d\lambda(x_j) \\
&= Z_A \cdot \overline{u_j} w_j \int_{\mathbb{F}} e^{-\nu_j |x_j|^2/2} \left|x_j\right|^2 \; d\lambda(x_j) \Big/ \int_{\mathbb{F}} e^{-\nu_j |x_j|^2/2} \, d\lambda(x_j)
\end{aligned}$$

One computes the quotient of these two integrals to be $d/\nu_j$, where as before $d = 1$ if $\mathbb{F} = \mathbb{R}$ and $d = 2$ if $\mathbb{F} = \mathbb{C}$. Putting things together gives

$$\begin{aligned}
\int_{\mathbb{F}^n} e^{-\frac{1}{2}\langle x, \tilde{A}x\rangle} \langle v_1, x\rangle \langle x, v_2\rangle \; d\lambda(x) &= d \cdot Z_A \sum_{j=1}^n \overline{u_j} w_j / \nu_j \\
&= dZ_A \left\langle Uv_1, D^{-1} Uv_2 \right\rangle \\
&= dZ_A \left\langle v_1, \tilde{A}^{-1} v_2 \right\rangle
\end{aligned}$$

□

## Problem 11.

Consider a sequence $(X_n)_{n\in\mathbb{N}}$ of independent and identically-distributed Bernoulli random variables, with a common parameter $p \in (0,1)$. Calculate the asymptotic distribution of the random variable $A_N := N^{-1}\sum_{n\leq N} X_n$ in two ways: by (a) appealing to the Central Limit Theorem and by (b) proving and then using the De Moivre-Laplace Theorem.

(a) Let $q := 1-p$. Each $X_n$ satisfies $\mathbb{P}[X_n = 0] = q$, $\mathbb{P}[X_n = 1] = p$. One calculates $\mathbb{E}[X_n] = p$, $\mathbb{E}[X_n^2] = p$ and so $\mathbb{V}\text{ar}[X_n] = pq$ and $\sigma_{X_n} = \sqrt{pq}$. By the Central Limit Theorem, we have $A_n \approx \mathcal{N}\left(np, \sqrt{pqn}\right)$

(b) Use the "De-Moivre Laplace Theorem" to recover the previous result.

*Solution.* An easy calculation shows that

$$\mathbb{P}_{A_n} = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \delta_k$$

The standard normal distribution with mean $pn$ and standard deviation $\sqrt{pqn}$ has probability density function

$$t \mapsto \frac{1}{\sqrt{2\pi pqn}} e^{-\frac{(t-np)^2}{2pqn}}$$

We would like to say that[9]

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi pqn}} e^{-\frac{(k-np)^2}{2pqn}} \tag{$*$}$$

The sense of $\approx$ will be clarified below. Stirling's approximation gives $x! \sim x^x e^{-x}\sqrt{2\pi x}$. It follows that

$$\begin{aligned}
\binom{n}{k} p^k q^{n-k} &= \frac{n!}{k!\cdot(n-k)!} p^k q^{n-k} \\
&\approx \frac{\sqrt{n}}{\sqrt{2\pi k(n-k)}} \cdot \frac{p^k q^{n-k}\cdot n^n/e^n}{k^k/e^k \cdot (n-k)^{n-k}/e^{n-k}} \\
&= \frac{\sqrt{n}}{\sqrt{2\pi k(n-k)}} \left[\left(\frac{pn}{k}\right)^{k/n} \left(\frac{qn}{n-k}\right)^{1-k/n}\right]^n \\
&= \frac{\sqrt{n}}{\sqrt{2\pi k(n-k)}} \exp\left(n\cdot[\eta\log\alpha + (1-\eta)\log\beta]\right)
\end{aligned}$$

where $\alpha := pn/k$, $\beta := qn/(n-k)$ and $\eta := k/n$. Since $\eta\alpha + (1-\eta)\beta = p + q = 1$, the concavity of the logarithm gives

$$\eta\log\alpha + (1-\eta)\log\beta \leq \log[\eta\alpha + (1-\eta)\beta] = 0$$

[9] To get $A_n \approx \mathcal{N}\left(np, \sqrt{pqn}\right)$, we should really show that $\binom{n}{k}p^k q^{n-k} \approx \int_{k-1/2}^{k+1/2} \frac{1}{\sqrt{2\pi pqn}} e^{-\frac{(t-np)^2}{2pqn}}\, dt$. But $\frac{1}{\sqrt{2\pi pqn}} e^{-\frac{(k-np)^2}{2pqn}}\, dt$ is a good proxy for this integral.

with equality if and only if $\alpha = \beta$.[10] It follows that $\binom{n}{k}p^k q^{n-k}$ tends exponentially to 0 if $\alpha \not\approx \beta$. Since $\eta\alpha + (1-\eta)\beta = 1$, we have $\alpha \approx \beta$ iff $\alpha \approx 1$, i.e. iff $p = k/n$. We will show that $(*)$ holds if $k/n \sim p$. Note that this implies $(n-k)/n \sim q$.

Going back to our Stirling-based approximation, and applying a Taylor expansion gives

$$\begin{aligned}
\eta \log \alpha + (1-\eta)\log\beta &= \frac{k}{n}\log\left(1 - \frac{k-pn}{k}\right) + \frac{n-k}{n}\log\left(1 + \frac{k-pn}{n-k}\right) \\
&\approx -\frac{k}{n}\left(\frac{k-pn}{k} + \frac{(k-pn)^2}{2k^2}\right) + \frac{n-k}{n}\left(\frac{k-pn}{n-k} - \frac{(k-pn)^2}{2(n-k)^2}\right) \\
&= -\frac{(k-pn)^2}{2nk} - \frac{(k-pn)^2}{2n(n-k)} \\
&\approx -\frac{(k-pn)^2}{2n^2p} - \frac{(k-pn)^2}{2n^2q} \\
&= -\frac{(k-pn)^2}{2n^2}\cdot\left(\frac{1}{p} + \frac{1}{q}\right) \\
&= -\frac{(k-pn)^2}{2n^2pq}
\end{aligned}$$

Thus

$$\begin{aligned}
\binom{n}{k}p^k q^{n-k} &\approx \frac{\sqrt{n}}{\sqrt{2\pi k(n-k)}}\exp\left(n\cdot[\eta\log\alpha + (1-\eta)\log\beta]\right) \\
&\approx \frac{\sqrt{n}}{\sqrt{2\pi pqn^2}}\exp\left(-n\cdot\frac{(k-pn)^2}{2n^2pq}\right) \\
&= \frac{1}{\sqrt{2\pi pqn}}\exp\left(-\frac{(k-pn)^2}{2npq}\right)
\end{aligned}$$

□

## Problem 12.

Let $Z$ be a standard normal random variable. Fixing $\mu \in \mathbb{R}$ and $\sigma > 0$, we define a new random variable

$$X := \exp(\mu + \sigma Z)$$

(a) Calculate the moments of $X$.

*Solution.* We have

$$\mathbb{E}[X^n] = \mathbb{E}[\exp(n\mu + n\sigma Z)] = \mathbb{E}[\exp(n\mu)\exp(n\sigma Z)] = e^{n\mu}\cdot\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{n\sigma t}\cdot e^{-t^2/2}\,dt$$

The integral at the end is given by

$$\int_{\mathbb{R}} e^{n\sigma t}\cdot e^{-t^2/2}\,dt = e^{n^2\sigma^2/2}\int_{\mathbb{R}} e^{-(t-n\sigma)^2/2}\,dt = e^{n^2\sigma^2/2}\int_{\mathbb{R}} e^{-t^2/2}\,dt = \sqrt{2\pi}e^{n^2\sigma^2/2}$$

Thus $\mathbb{E}[X^n] = e^{n\mu + n^2\sigma^2/2}$. □

[10] We also have equality if $\eta = 0$ or $\eta = 1$. The special cases $k \approx 0$ and $k \approx n$ therefore need to be checked separately.

(b) Show that $\mathbb{E}[e^{tX}] = \infty$ for all $t > 0$.

*Solution.* Note that $X$ takes only positive values. Therefore, by the monotone convergence theorem,

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right] = \sum_{n=0}^{\infty} \mathbb{E}\left[\frac{t^n X^n}{n!}\right] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n] = \sum_{n=0}^{\infty} \frac{t^n}{n!} e^{n\mu + n^2\sigma^2/2}$$

If this value were finite, the summands would tend to 0. However, the bound $n! \le n^n$ shows that

$$\frac{t^n}{n!} e^{n\mu+n^2\sigma^2/2} \ge \frac{t^n}{n^n} e^{n\mu+n^2\sigma^2/2} = e^{n\mu+n^2\sigma^2/2+n\log t - n\log n}$$

Since $\sigma > 0$, the exponent

$$n\mu + n^2\sigma^2/2 + n\log t - n\log n$$

is dominated by the term $n^2\sigma^2/2$, so tends to $\infty$ with $n$. It follows that the series diverges to infinity. So $\mathbb{E}[e^{tX}] = \infty$. □

(c) Show that $Z$ has the same moments as some *discrete* random variable on $\mathbb{R}$ (i.e. one whose measure is supported on a countable set).

*Solution.* Let $\xi$ be the measure

$$\xi := C \sum_{m\in\mathbb{Z}} e^{-\sigma^2 m^2/2} \delta_{\sigma m}$$

where $C = \left(\sum_{m\in\mathbb{Z}} e^{-\sigma^2 m^2/2}\right)^{-1}$ is chosen so that $\xi(\mathbb{R}) = 1$. Let $\nu$ be the pushforward of $\xi$ under the map $x \mapsto \exp(\mu + \sigma x)$. Concretely, we have

$$\nu := C \sum_{m\in\mathbb{Z}} e^{-\sigma^2 m^2/2} \delta_{e^{\mu+\sigma^2 m}}$$

Then the $n$-th moment of $\nu$ is given by

$$\begin{aligned}\int_{\mathbb{R}} x^n \, d\nu(x) &= C \sum_{m\in\mathbb{Z}} e^{n(\mu+\sigma^2 m)} e^{-\sigma^2 m^2/2} \\ &= e^{n\mu+\sigma^2 n^2/2} \cdot C \sum_{m\in\mathbb{Z}} e^{-\sigma^2 (m-n)^2/2} \\ &= e^{n\mu+\sigma^2 n^2/2} \cdot C \sum_{m\in\mathbb{Z}} e^{-\sigma^2 m^2/2} \\ &= e^{n\mu+\sigma^2 n^2/2} = \mathbb{E}[X^n]\end{aligned}$$

□