Homework 2

SDS 391P.6, Spring 2026

Pratik Patil

Due: Feb 16 (Monday)

This homework is (also) a work in progress and is provided as-is for instructional purposes. The problems are drawn from various sources and do not (yet) have sufficient references to the original material. Additionally, errors may be present. so some cauution is advised! The document will be updated if corrections are necessary. Last updated: 2026-02-02.

0 Guidelines

- Please start early. If you have questions about the statements, notation, or possible typos, email us as soon as possible. When emailing about the course, please begin the subject line with [SDS 391P.6].
- Please begin your answer to each *main* question on a separate page. If you use any code, include it in an appendix. Submit a single combined PDF to Canvas. (If you encounter any submission issues, please let us know.)
- The problems and motivations draw on multiple sources (past course material, textbooks, and other standard references). If you use any external resources that materially guide your solution (beyond routine lookups), please cite them in your write-up.
- These questions are designed to build intuition and technique. You are welcome to go beyond what is explicitly asked. If you introduce additional assumptions (while keeping the spirit of the problem), state them clearly. If you discover something interesting along the way, feel free to include it as a brief remark; we may share especially instructive observations with the class.
- Parts labeled [Bonus] are optional: they are not required for full credit. They are intended as extra practice. You may skip them without penalty. If you attempt them, please label your solutions clearly with [Bonus].
- Many parts include hints intended to help you get started. You are not required to follow the suggested route, and you are encouraged to try alternative approaches when appropriate.
- We will grade primarily for correctness and clear reasoning. Do not over-optimize for minor presentation details. The spirit of the homework is for you to learn something new!

1 Jackknife: variance and bias

This exercise is meant to help practice variance tensorization and Efron–Stein–Steele (ESS) style arguments in the context of the classical *jackknife* (leave-one-out) setting (Quenouille, 1949). (If you are curious about the history and the name "jackknife", please see Section 7 of Cosma Shalizi's notes: `https://www.stat.cmu.edu/~cshalizi/490/23/2023-11-29-jackknife.pdf`)

To describe the setting, let $X_1, \ldots, X_n$ be i.i.d. real-valued random variables. (Assume throughout that all random variables below have finite second moments whenever a variance is invoked.) Let θ be a parameter of the distribution of the X_i's that we wish to estimate. For each $k \in \mathbb{N}$, let $f_k : \mathbb{R}^k \to \mathbb{R}$ be a (measurable) statistic that is *symmetric* (invariant under permutations of its arguments).

Define the full-sample estimator

$$Z_n := f_n(X_1, \ldots, X_n).$$

The quality of the estimator Z_n is measured by by its bias and its variance (to be defined below shortly). Since we do not know the distribution of the X_i's, we need to estimate the bias and variance. The key difficulty is that we have only *one* dataset and so the original estimator Z_n and the bias and variance estimators are to be constructed from the *same* data. The jackknife is a technique to (partially) solve this difficulty.

Define the leave-one-out (jackknife) replicates

$$Z^{(i)} \;:=\; f_{n-1}(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n), \qquad i = 1, \ldots, n,$$

and their average

$$\overline{Z} \;:=\; \frac{1}{n} \sum_{i=1}^{n} Z^{(i)}.$$

Below we will define jackknife estimates of the variance and bias and show some of their properties.

(a) **Variance bound.** Define the (simplified) jackknife variance functional

$$\widehat{V} \;:=\; \sum_{i=1}^{n} \big(Z^{(i)} - \overline{Z}\big)^2.$$

(a1) Show the identity

$$\widehat{V} \;=\; \sum_{i=1}^{n} (Z^{(i)} - \overline{Z})^2 \;=\; \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{n} \big(Z^{(i)} - Z^{(j)}\big)^2.$$

(*Hint:* Start from $\sum_i (Z^{(i)} - \overline{Z})^2 \;=\; \sum_i (Z^{(i)})^2 - n\overline{Z}^2$. Then expand the double sum $\sum_{i,j} (Z^{(i)} - Z^{(j)})^2$ and compare terms.)

(a2) Let

$$W \;:=\; f_{n-1}(X_1, \ldots, X_{n-1})$$

be the same statistic applied to an i.i.d. sample of size $n-1$. Show that

$$\mathrm{Var}(W) \;\leqslant\; \mathbb{E}[\widehat{V}].$$

In words: $\widehat{V}$ *overestimates* the variance of f_{n-1} on average.

(*Hint:* By i.i.d. symmetry, $\mathrm{Var}(W) \;=\; \mathrm{Var}(Z^{(1)})$. Apply tensorization of variance to $Z^{(1)} = f_{n-1}(X_2, \ldots, X_n)$ as a function of the independent inputs $X_2, \ldots, X_n$:

$$\mathrm{Var}(Z^{(1)}) \;\leqslant\; \sum_{k=2}^{n} \mathbb{E}\Big[\mathrm{Var}\Big(Z^{(1)} \,\Big|\, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n\Big)\Big].$$

Fix $k \in \{2, \ldots, n\}$ and condition on $(X_2, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n)$. Using symmetry of f_{n-1}, write $Z^{(1)} = g(X_k)$ and $Z^{(k)} = g(X_1)$ for the *same* (random) function g. Under the conditioning, X_1 and X_k are i.i.d. and independent of the conditioning, hence

$$\mathrm{Var}(Z^{(1)} \mid \cdots) = \mathrm{Var}(g(X_k) \mid \cdots) = \frac{1}{2}\,\mathbb{E}\big[(g(X_k) - g(X_1))^2 \mid \cdots\big] = \frac{1}{2}\,\mathbb{E}\Big[(Z^{(1)} - Z^{(k)})^2 \mid \cdots\Big].$$

Sum over $k = 2, \ldots, n$ and take expectations to obtain

$$\mathrm{Var}(Z^{(1)}) \;\leqslant\; \frac{1}{2}\sum_{k=2}^{n} \mathbb{E}\big[(Z^{(1)} - Z^{(k)})^2\big].$$

Finally, use exchangeability of $(Z^{(1)}, \ldots, Z^{(n)})$ together with part (a1) to relate the right-hand side to $\mathbb{E}[\widehat{V}]$.)

(b) **[Bonus] Exact bias for quadratic statistics.** Let θ be a parameter of the distribution of X_1, and define the full-sample estimator

$$Z_n := f_n(X_1, \ldots, X_n).$$

Assume the bias admits a first-order expansion

$$\mathbb{E}Z_n - \theta \;=\; \frac{c}{n} + O(n^{-2}) \qquad \text{as } n \to \infty,$$

for some constants $c \in \mathbb{R}$.

Define the jackknife bias estimate

$$\widehat{B}_n \;:=\; (n-1)\big(\overline{Z} - Z_n\big),$$

and the bias-corrected estimator

$$\widetilde{Z}_n \;:=\; Z_n - \widehat{B}_n \;=\; nZ_n - (n-1)\overline{Z}.$$

Show that

$$\mathbb{E}\widetilde{Z}_n - \theta \;=\; O(n^{-2}).$$

Thus, for quadratic statistics (which can be written as degree two polynomials) with bias of order $O(n^{-1})$, the jackknife can provide exact bias!

(*Hint:* First show $\mathbb{E}[\overline{Z}] = \mathbb{E}[f_{n-1}(X_1, \ldots, X_{n-1})] = \mathbb{E}Z_{n-1}$ by exchangeability. Then compute

$$\mathbb{E}\widetilde{Z}_n \;=\; n\,\mathbb{E}Z_n - (n-1)\,\mathbb{E}Z_{n-1}$$

and plug in the bias expansions for n and $n-1$.)

2 Order statistics: variance versus spacings

This exercise is meant to provide practice with the *one-sided* ESS inequality to relate the variance of an order statistic to the squared size of a neighboring spacing.

Let $X_1, \ldots, X_n$ be independent real-valued random variables. Let

$$X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$$

denote their order statistics, and define the spacings

$$\Delta_k := X_{(k+1)} - X_{(k)}, \qquad k = 1, \dots, n-1.$$

Assume $\mathbb{E}[X_{(k)}^2] < \infty$ for the relevant k so that the variances below are finite.

For each $i \in \{1, \dots, n\}$, let X_i' be an independent copy of X_i (independent of everything), and let $(X_1^{(i)}, \dots, X_n^{(i)})$ be the resampled vector where

$$X_j^{(i)} := \begin{cases} X_j, & j \neq i, \\ X_i', & j = i. \end{cases}$$

Let $X_{(k)}^{(i)}$ denote the kth order statistic of the resampled sample.

Recall the one-sided ESS inequality:

$$\mathrm{Var}(Z) \;\leqslant\; \sum_{i=1}^{n} \mathbb{E}\big[(Z - Z^{(i)})_+^2\big], \qquad (a)_+ := \max\{a, 0\},$$

as well as the analogous bound with $(Z^{(i)} - Z)_+^2$ (apply the inequality to $-Z$). Throughout, we use $\mathbb{1}\{E\}$ for indicator functions.

(a) **The maximum.** We will first work with the extreme order statistic, the maximum, and obtain a bound on its variance. We will then check how good this bound is for two concrete distributions.

(a1) Show that, regardless of the distributions of the independent X_i,

$$\mathrm{Var}\big(X_{(n)}\big) \;\leqslant\; \mathbb{E}\big[(X_{(n)} - X_{(n-1)})^2\big] \;=\; \mathbb{E}[\Delta_{n-1}^2].$$

(*Hint:* Prove the pointwise bound

$$(X_{(n)} - X_{(n)}^{(i)})_+ \;\leqslant\; (X_{(n)} - X_{(n-1)})\, \mathbb{1}\{X_i = X_{(n)}\}.$$

Then sum over i and apply one-sided ESS. If you worry about ties: when $X_{(n)} = X_{(n-1)}$, the spacing is 0 and the bound is automatic. You may also assume the X_i's are a.s. distinct to avoid worrying about ties.)

(a2) Let us check the bound for a uniform distribution. Assume $X_1, \dots, X_n$ are i.i.d. Unif$[0,1]$. Compute explicitly:

$$\mathrm{Var}(X_{(n)}), \qquad \mathbb{E}\big[(X_{(n)} - X_{(n-1)})^2\big].$$

Also compute the limits as $n \to \infty$ and the ratio of the right-hand side to the left-hand side.

(*Hint:* $X_{(n)} \sim \mathrm{Beta}(n,1)$. Also, $\Delta_{n-1} = X_{(n)} - X_{(n-1)} \sim \mathrm{Beta}(1,n)$.)

(a3) [Bonus] Let us now check the bound for an exponential distribution. Assume $X_1, \dots, X_n$ are i.i.d. Exp(1) with density $e^{-x}\, \mathbb{1}\{x \geqslant 0\}$. Compute explicitly:

$$\mathrm{Var}(X_{(n)}), \qquad \mathbb{E}\big[(X_{(n)} - X_{(n-1)})^2\big].$$

Also compute the limits as $n \to \infty$ and the ratio of the right-hand side to the left-hand side.

(b) **[Bonus] General order statistics.** We now generalize the previous result for any order statistic. Fix $k \in \{2, \ldots, n-1\}$.

(b1) Prove the two inequalities

$$\operatorname{Var}\big(X_{(k)}\big) \;\leqslant\; k\,\mathbb{E}\big[(X_{(k+1)} - X_{(k)})^2\big] = k\,\mathbb{E}[\Delta_k^2],$$

and

$$\operatorname{Var}\big(X_{(k)}\big) \;\leqslant\; (n-k+1)\,\mathbb{E}\big[(X_{(k)} - X_{(k-1)})^2\big] = (n-k+1)\,\mathbb{E}[\Delta_{k-1}^2].$$

(At the endpoints, interpret the first inequality for $k = 1$ and the second for $k = n$.)

(*Hint:* Apply one-sided ESS to $Z = X_{(k)}$ and also to $-Z = -X_{(k)}$. A useful goal is to show the pointwise comparisons

$$(X_{(k)}^{(i)} - X_{(k)})_+ \;\leqslant\; (X_{(k+1)} - X_{(k)})\,\mathbb{1}\{X_i \leqslant X_{(k)}\},$$

and

$$(X_{(k)} - X_{(k)}^{(i)})_+ \;\leqslant\; (X_{(k)} - X_{(k-1)})\,\mathbb{1}\{X_i \geqslant X_{(k)}\}.$$

Then use the counting bounds $\sum_{i=1}^n \mathbb{1}\{X_i \leqslant X_{(k)}\} \leqslant k$ and $\sum_{i=1}^n \mathbb{1}\{X_i \geqslant X_{(k)}\} \leqslant n-k+1$. Finally, as for (a1), ties can be handled, but you may also assume the X_i's are a.s. distinct to simplify.)

(b2) Deduce the "best of the two" corollary:

$$\operatorname{Var}(X_{(k)}) \;\leqslant\; \begin{cases} k\,\mathbb{E}[\Delta_k^2], & 1 \leqslant k \leqslant \lfloor n/2 \rfloor, \\ (n-k+1)\,\mathbb{E}[\Delta_{k-1}^2], & \lfloor n/2 \rfloor < k \leqslant n. \end{cases}$$

3 Rademacher processes: bounding the variance

This exercise studies a basic Rademacher process and bounds variance of its suprema. We will study expectations of suprema of random processes during the second part of the course.

Let $\varepsilon_1, \ldots, \varepsilon_n$ be independent Rademacher random variables, $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$. Let $T \subset \mathbb{R}^n$ be a finite set, and write $t = (t_1, \ldots, t_n)$ for $t \in T$. Define

$$Z \;:=\; \sup_{t \in T} \sum_{i=1}^n \varepsilon_i t_i.$$

(As T is finite, the supremum is a maximum.)

Introduce the parameters

$$\sigma^2 \;:=\; \sup_{t \in T} \sum_{i=1}^n t_i^2 \qquad \text{and} \qquad \sigma_\infty^2 \;:=\; \sum_{i=1}^n \sup_{t \in T} t_i^2,$$

and write $(a)_+ := \max\{a, 0\}$.

(a) **A weak variance bound.** Let us start simple and show a coordinatewise variance bound for Z.

(a1) For fixed $t \in \mathbb{R}^n$, show that

$$\operatorname{Var}\left(\sum_{i=1}^n \varepsilon_i t_i\right) \;=\; \sum_{i=1}^n t_i^2.$$

Conclude that

$$\sup_{t\in T} \operatorname{Var}\left(\sum_{i=1}^n \varepsilon_i t_i\right) \;=\; \sigma^2.$$

(*Hint:* Use $\mathbb{E}[\varepsilon_i] = 0$, $\mathbb{E}[\varepsilon_i^2] = 1$, and independence to kill cross terms.)

(a2) Show that

$$\operatorname{Var}(Z) \;\leqslant\; \sigma_\infty^2.$$

(*Hint:* Use the bounded differences variance bound: if flipping ε_i changes Z by at most c_i a.s., then $\operatorname{Var}(Z) \leqslant \frac{1}{4}\sum_i c_i^2$. Here $c_i = 2\sup_{t\in T}|t_i|$, so $\operatorname{Var}(Z) \leqslant \sum_i \sup_{t\in T} t_i^2 = \sigma_\infty^2$.)

(b) **[Bonus] A sharper variance bound.** Now we show that the supremum inside variance costs only a constant factor.

(b1) Prove that

$$\operatorname{Var}(Z) \;\leqslant\; 2\sigma^2.$$

(*Hint:* Let $\varepsilon_1', \ldots, \varepsilon_n'$ be an independent copy of $\varepsilon_1, \ldots, \varepsilon_n$ and define

$$\varepsilon^{(i)} := (\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_i', \varepsilon_{i+1}, \ldots, \varepsilon_n), \qquad Z^{(i)} := \sup_{t\in T} \sum_{j=1}^n \varepsilon_j^{(i)} t_j.$$

Fix a deterministic tie-breaking rule (e.g., pick the smallest index in a fixed ordering of T) and let $t^* = t^*(\varepsilon) \in T$ be a measurable maximizer so that $Z = \sum_{j=1}^n \varepsilon_j t_j^*$. Show that

$$(Z - Z^{(i)})_+ \;\leqslant\; |\varepsilon_i - \varepsilon_i'|\,|t_i^*|.$$

Then compute $\mathbb{E}[(\varepsilon_i - \varepsilon_i')^2] = 2$ and use one-sided ESS: $\operatorname{Var}(Z) \leqslant \sum_{i=1}^n \mathbb{E}[(Z - Z^{(i)})_+^2]$. Finally, use $\sum_i (t_i^*)^2 \leqslant \sigma^2$ pointwise since $t^* \in T$.)

(b2) Give an example of a finite $T \subset \mathbb{R}^n$ for which

$$\frac{\sigma_\infty^2}{\sigma^2} \asymp n,$$

so the bound in part (b1) improves the bound in part (a2) by a factor of order n.

(*Hint:* Take $T = \{e_1, \ldots, e_n\}$, the standard basis vectors.)

4 Polynomial versus exponential moment method bounds

This exercise compares two ways of optimizing Markov's inequality: using polynomial moments or using exponential moments (Chernoff bounds). You will show that polynomial-moment optimization is *always* at least as good numerically, but is often much harder to use.

Let Y be a nonnegative random variable and fix $t > 0$. Define

$$M(t) \;:=\; \inf_{q\in\mathbb{Z}_+} \frac{\mathbb{E}[Y^q]}{t^q}, \qquad \text{(with the convention } Y^0 \equiv 1\text{)},$$

and

$$C(t) \;:=\; \inf_{\lambda>0} \mathbb{E}\Big[e^{\lambda(Y-t)}\Big] \;=\; \inf_{\lambda>0} e^{-\lambda t}\,\mathbb{E}e^{\lambda Y}.$$

(Here $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$.)

Recall from the lecture notes that both $M(t)$ and $C(t)$ arise by applying Markov's inequality to Y^q and $e^{\lambda Y}$, respectively, so $\mathbb{P}\{Y \geqslant t\} \leqslant M(t)$ and $\mathbb{P}\{Y \geqslant t\} \leqslant C(t)$.

(a) **Polynomial moments are at least as good.** Prove that

$$M(t) \;\leqslant\; C(t) \qquad \text{for every } t > 0.$$

(*Hint:* Fix $\lambda > 0$ and use the power series $e^{\lambda y} = \sum_{q=0}^{\infty} \frac{\lambda^q y^q}{q!}$ for $y \geqslant 0$. By monotone convergence,

$$e^{-\lambda t}\mathbb{E}e^{\lambda Y} = \sum_{q=0}^{\infty} w_q(\lambda, t)\,\frac{\mathbb{E}[Y^q]}{t^q}, \qquad w_q(\lambda, t) := e^{-\lambda t}\frac{(\lambda t)^q}{q!}.$$

Note that $(w_q(\lambda, t))_{q\geqslant 0}$ is a probability distribution (Poisson(λt) PMF), so the right-hand side is a convex combination of the values $\mathbb{E}[Y^q]/t^q$ and is therefore at least their minimum. Then take $\inf_{\lambda>0}$.)

(b) **Chernoff is still useful.** Give a brief explanation (a few sentences) of why Chernoff bounds are nevertheless widely used in practice, especially for sums of independent random variables.

5 Sub-Gaussian characterizations

This exercise proves several equivalent (up to constants) ways of expressing that a centered random variable has Gaussian-type tails. The three most common viewpoints are: (i) a quadratic CGF bound, (ii) a Gaussian tail bound, (iii) Gaussian-like moment growth.

Let Z be a real-valued random variable with $\mathbb{E}Z = 0$. Define the centered CGF

$$\psi_Z(\lambda) \;:=\; \log \mathbb{E}e^{\lambda Z} \in (-\infty, \infty] \qquad (\lambda \in \mathbb{R}).$$

(a) **Three sub-Gaussian properties.** Fix $v \geqslant 0$. Consider the statements:

(i) CGF bound: For all $\lambda \in \mathbb{R}$,

$$\psi_Z(\lambda) \;\leqslant\; \frac{v\lambda^2}{2}.$$

(ii) Tail bound: For all $t \geqslant 0$,

$$\mathbb{P}\{|Z| \geqslant t\} \;\leqslant\; 2\exp\Big(-\frac{t^2}{2v}\Big).$$

(iii) Moment growth: There exists an absolute constant $C > 0$ such that for all $p \geqslant 1$,

$$\big(\mathbb{E}|Z|^p\big)^{1/p} \;\leqslant\; C\sqrt{vp}.$$

Our goal is to prove that (i), (ii), (iii) are equivalent up to universal constants by establishing the implication chain

$$\text{(i)} \;\Rightarrow\; \text{(ii)} \;\Rightarrow\; \text{(iii)} \;\Rightarrow\; \text{(i')}$$

where (i') is the same as (i) but with v replaced by $c \cdot v$ for a universal constant $c > 0$.

(a1) (i) $\Rightarrow$ (ii). Use the Chernoff bound (Laplace transform method) on Z and on $-Z$.

(*Hint:* For $\lambda \geqslant 0$, $\mathbb{P}\{Z \geqslant t\} \leqslant \exp(-\lambda t + \psi_Z(\lambda))$. Insert $\psi_Z(\lambda) \leqslant v\lambda^2/2$ and optimize in λ; then repeat for $-Z$.)

(a2) (ii) $\Rightarrow$ (iii). Recall from Homework 1, Question 4, Part (a1) that for $p > 0$,

$$\mathbb{E}|Z|^p \;=\; p\int_0^\infty t^{p-1}\,\mathbb{P}\{|Z| \geqslant t\}\,dt.$$

Use this tail bound to estimate the integral. (Technically, you can also use Homework 1, Question 4, Part (a2) directly.)

(*Hints:* Similar to Homework 1, Question 4, Part (a2), you may use the fact that $\Gamma(s) = \int_0^\infty u^{s-1}e^{-u}\,du \leqslant s^s$ for $s \geqslant 1$.)

(a3) [Bonus] (iii) $\Rightarrow$ (i'). Assume the moment growth bound in (iii). Show that there is a universal constant $c > 0$ (depending only on C in (iii)) such that for all $\lambda \in \mathbb{R}$,

$$\log \mathbb{E}e^{\lambda Z} \;\leqslant\; \frac{(cv)\lambda^2}{2}.$$

(*Hint:* Start from

$$\mathbb{E}e^{\lambda Z} \;=\; \sum_{k=0}^{\infty} \frac{\lambda^k\,\mathbb{E}[Z^k]}{k!}, \qquad |\mathbb{E}[Z^k]| \leqslant \mathbb{E}|Z|^k.$$

Use (iii) to bound $\mathbb{E}|Z|^k$, then compare to $\exp(\text{const}\cdot v\lambda^2)$ using Stirling's bound $k! \geqslant (k/e)^k$. A good strategy is to split the series into "small k" and "large k" relative to $v\lambda^2$.)

(b) **Variance proxy.** Assume the CGF bound (i) holds for some v. Show that $\mathrm{Var}(Z) \leqslant v$.

(*Hint:* Use the fact discussed in lecture that whenever $\mathbb{E}e^{\lambda Z} < \infty$ near 0, we have $\psi_Z''(0) = \mathrm{Var}(Z)$. Differentiate the inequality $\psi_Z(\lambda) \leqslant v\lambda^2/2$ twice at $\lambda = 0$.)

(c) **[Bonus] Exponential-square integrability.** Assume the tail bound (ii). Show that there is a universal constant $c > 0$ such that

$$\mathbb{E}\exp\!\Big(\frac{Z^2}{c\,v}\Big) \;\leqslant\; 2.$$

Conversely, show that if $\mathbb{E}\exp(Z^2/(c\,v)) \leqslant 2$ for some universal constant $c > 0$, then Z is sub-Gaussian in the sense of (i') for a (possibly different) universal constant.

(*Hints:* For the first direction, use the tail integral identity from (c2) with $\Phi(t) = \exp(t^2/(cv))$. For the second direction, use Young's inequality

$$|\lambda z| \leqslant \frac{a\lambda^2}{2} + \frac{z^2}{2a} \quad \text{for any } a > 0,$$

then take expectations and choose a as a constant multiple of v.)

Source material

Parts of this homework were inspired by exercises from Boucheron et al. (2013); Tropp (2023); van Handel (2016), in addition to the author's accumulated experience working on related topics.

References

Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration Inequalities - A Nonasymptotic Theory of Independence*. Oxford University Press.

Quenouille, M. H. (1949). Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 11(1):68–84.

Tropp, J. A. (2023). Probability in high dimensions. Caltech CMS Lecture Notes 2021-01.

van Handel, R. (2016). Probability in high dimension. Lecture Notes (Princeton University). APC 550.