1.1 ☕ Consider any subset $T \subset \mathbb{R}^n$. Check that $\mathrm{conv}(T)$ is a convex set.

A CONVEX COMBINATION OF POINTS $x_1, \dots, x_m \in \mathbb{R}^n$ IS A LINEAR COMBINATION WITH NONNEGATIVE COEFFICIENTS THAT ADD TO 1:

$$\sum_{i=1}^{m} \lambda_i x_i, \quad \sum_{i=1}^{m} \lambda_i = 1, \quad \lambda_i \geq 0 \text{ FOR ALL } i.$$

THE CONVEX HULL OF A SET $T \subseteq \mathbb{R}^n$ IS THE SET OF ALL CONVEX COMBINATIONS OF ALL FINITE COLLECTIONS OF POINTS IN T:

$$\mathrm{conv}(T) = \{\text{CONVEX COMBINATIONS OF } x_1, \dots, x_m \in T, \ m \in \mathbb{N}\}.$$

$\mathrm{conv}(T)$ IS CONVEX IF

$$\lambda x + (1-\lambda) y \in \mathrm{conv}(T), \quad \forall x, y \in \mathrm{conv}(T),$$

WHICH IS TRUE BY DEF. OF THE CONVEX HULL.

1.2 ☕ Check that the pointwise maximum of a finite number of convex functions is a convex function.

LET K BE A CONVEX SET. A FUNCTION $f: K \to \mathbb{R}$ IS CONVEX IF

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in K, \ \lambda \in [0,1].$$

LET $f_i : K \to \mathbb{R}$, WITH $i = 1, \dots, n$, BE A COLLECTION OF CONVEX FUNCTIONS. DEFINE THE POINTWISE MAXIMUM FUNCTION AS

$$g: K \to \mathbb{R}$$
$$x \mapsto \max\{f_1(x), \dots, f_n(x)\}.$$

THEN

$$f_i(\lambda x + (1-\lambda) y) \leq \lambda f_i(x) + (1-\lambda) f_i(y)$$
$$\leq \lambda g(x) + (1-\lambda) g(y),$$

FOR $i = 1, \dots, n$. THEREFORE,

$$\max_i f_i(\lambda x + (1-\lambda)y) \le \lambda g(x) + (1-\lambda) g(y),$$

WHICH PROVES THE RESULT.

1.3 ☕☕ (Jensen inequality)

(a) The definition of a convex function (1.1) involves convex combinations of two points x and y. Let us extend it to arbitrarily many points. Let $K \subset \mathbb{R}^n$ be a convex subset. Prove that a function $f : K \to \mathbb{R}$ is convex if and only if the following holds. For any $m \in \mathbb{N}$, any vectors $x_i \in K$ and any numbers $\lambda_i \ge 0$ with $\sum_{i=1}^m \lambda_i = 1$, we have

$$f\Big(\sum_{i=1}^m \lambda_i x_i\Big) \le \sum_{i=1}^m \lambda_i f(x_i).$$

(b) Let X be a random vector in $\mathbb{R}^n$ that takes finitely many values, and let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Deduce from part (a) Jensen inequality:

$$f(\mathbb{E}\, X) \le \mathbb{E}\, f(X).$$

(a) "$\Leftarrow$" TRIVIAL.

"$\Rightarrow$" AS $f: K \to \mathbb{R}$ IS CONVEX,

$$f(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 f(x) + \lambda_2 f(y), \quad \forall x, y \in K, \ \lambda_1, \lambda_2 \in [0,1], \ \lambda_1 + \lambda_2 = 1.$$

THEREFORE, THE IMPLICATION HOLDS FOR $m = 2$, WHICH SUGGEST THAT WE CAN PROCEED BY INDUCTION. THEREFORE, LET'S ASSUME THAT THE IMPLICATION HOLDS FOR $m = n$. THEN, AS K IS CONVEX, $\sum_{i=1}^{n+1} \lambda_i x_i \in K$ FOR ALL $x_1, \dots, x_{n+1} \in K$ AND $\lambda_1, \dots, \lambda_{n+1} \in [0,1]$ WITH $\lambda_1 + \dots + \lambda_{n+1} = 1$. LET $\lambda_n^* = \sum_{i=1}^n \lambda_i$. THEN $\sum_{i=1}^n \frac{\lambda_i}{\lambda_n^*} x_i \in K$ BY THE CONVEXITY OF K AND

$$f\Big(\sum_{i=1}^{n+1} \lambda_i x_i\Big) = f\Big(\sum_{i=1}^n \lambda_i x_i + \lambda_{n+1} x_{n+1}\Big)$$

$$= f\Big(\lambda_n^* \underbrace{\sum_{i=1}^n \frac{\lambda_i}{\lambda_n^*} x_i}_{\in K} + \lambda_{n+1} x_{n+1}\Big)$$

$$\le \lambda_n^* f\Big(\sum_{i=1}^n \frac{\lambda_i}{\lambda_n^*} x_i\Big) + \lambda_{n+1} f(x_{n+1}) \quad \text{(BY THE CONVEXITY OF } f\text{)}.$$

$$\geq \cancel{\lambda_n^*} \sum_{i=1}^{n} \frac{\lambda_i}{\cancel{\lambda_n^*}} f(x_i) + \lambda_{n+1} f(x_{n+1}) \quad \text{(BY INDUCTION HYP.)}$$

$$= \sum_{i=1}^{n+1} \lambda_i f(x_i).$$

(b) ASSUME THAT $X \in \{x_1, \dots, x_m\}$ WITH RESPECTIVE PROBABILITIES $\{\lambda_1, \dots, \lambda_m\}$. THEN

$$\mathbb{E} X = \sum_{i=1}^{m} \lambda_i x_i,$$

AND PART (a) IMPLIES THAT

$$f(\mathbb{E} X) = f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i) = \mathbb{E} f(X).$$

1.4 ☕ (A maximum principle) Prove for any convex function f and a subset $T \subset \mathbb{R}^n$:

$$\sup_{x \in \mathrm{conv}(T)} f(x) = \sup_{x \in T} f(x).$$

BY DEF, $T \subseteq \mathrm{CONV}(T)$. THEREFORE,

$$\sup_{x \in \mathrm{conv}(T)} f(x) \geq \sup_{x \in T} f(x).$$

TO PROVE THE OPPOSITE DIRECTION, OBSERVE THAT, AS f IS CONVEX, BY CARATHÉODORY'S TH., FOR EVERY $x \in \mathrm{CONV}(T)$, THERE ARE $x_1, \dots, x_{n+1} \in T$ AND $\lambda_1, \dots, \lambda_{n+1} \in [0,1]$ WITH $\sum_{i=1}^{n+1} \lambda_i = 1$ S.T. $x = \sum_{i=1}^{n+1} \lambda_i x_i$. THEREFORE, BY THE CONVEXITY OF f,

$$f(x) = f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leq \sum_{i=1}^{n+1} \lambda_i f(x_i) \leq \sup_{x \in T} \sum_{i=1}^{n+1} \lambda_i f(x) = \sup_{x \in T} f(x)$$

WHERE THE FIRST INEQ. IS OBTAINED BY THE DEF. OF CONVEXITY OF f EXERCISE 1.3.a. THEN IMPLIES THAT

$$\sup_{x \in \operatorname{conv}(T)} f(x) \leq \sup_{x \in T} f(x).$$

1.5 ☕☕ (Expressing a cube as a convex hull of its vertices) It seems almost obvious that the cube is the convex hull of its vertices:

$$[-1,1]^n = \operatorname{conv}\left(\{-1,1\}^n\right).$$

Prove this by expressing any point in the cube as a convex combination of the vertices.

LET'S CALL $T := \{-1,1\}^n$ AND $X := \operatorname{conv}(T)$. LET $x = (x_1, \ldots, x_n) \in X$. THEN $x_i \in [-1,1]$. LET $v = (v_1, \ldots, v_n) \in T$. THEN $v_i \in \{-1,1\}$. WE WANT x_i TO BE A CONVEX COMB. OF 1 AND -1: FOR $p_i \in [0,1]$,

$$\begin{aligned} x_i &= p_i \cdot 1 + (1-p_i)(-1) \\ &= 2p_i - 1. \end{aligned}$$

THEN $p_i = (x_i+1)/2 \in [0,1]$. DEFINE $\lambda_v = \prod_{i=1}^{n} \alpha(v_i)$, WHERE

$$\alpha(v_i) = \begin{cases} p_i & \text{IF } v_i = 1, \\ 1-p_i & \text{IF } v_i = -1. \end{cases}$$

THINK OF λ_v AS THE DISTRIB. OF A RANDOM VECTOR $V = (V_1, \ldots, V_n)$ WITH INDEP. ENTRIES, WHERE $P[V_i = 1] = p_i$ AND $P[V_i = -1] = 1 - p_i$. THEN

$$P[V = v] = \prod_{i=1}^{n} P[V_i = v_i] = \lambda_v.$$

THEREFORE,

$$1 = \sum_{v \in T} P[V = v] = \sum_{v \in T} \lambda_v$$

SHOWING THAT $\lambda_1, \ldots \lambda_n$ ARE CONVEX COEFFICIENTS. EXPLICITLY,

$$\begin{aligned} \sum_{v \in T} \lambda_v &= \sum_{v \in T} \prod_{i=1}^{n} \alpha(v_i) \\ &= \prod_{i=1}^{n} \left[p_i + (1-p_i)\right] \end{aligned}$$

(THINK ABOUT EXPANDING $\prod_{i=1}^{n}(a_i + b_i)$)

$= 1.$

ALSO,

$$\sum_{v \in T} \lambda_v \cdot v = \mathbb{E}V = (\mathbb{E}V_1, \ldots, \mathbb{E}V_n) = (2p_1 - 1, \ldots, 2p_n - 1) = (x_1, \ldots, x_n) = x,$$

WHICH PROVES THAT x IS A CONVEX COMBINATION OF ALL $v \in T$.

1.6 ☕☕ (Expressing a cross-polytope as a convex hull of its vertices) Check that the unit ball corresponding to the ℓ^1 norm in $\mathbb{R}^n$ is the *absolute convex hull* of the standard basis $e_1, \ldots, e_n$ in $\mathbb{R}^n$, that is

$$B_1^n = \operatorname{conv}(\{\pm e_1, \ldots, \pm e_n\}).$$

Write down a formula that expresses any point $x \in B_1^n$ as a convex combination of the vectors $\pm e_1, \ldots, \pm e_n$.

FOR $x = (x_1, \ldots, x_n) \in B_1^n = \{x \in \mathbb{R}^n : \|x\|_1 \leq 1\}$, $\|x\|_1 = r \leq 1$. SO LET $s = 1 - r$.

CALL $\mu_i = \max\{x_i, 0\}$ AND $\mu_{-i} = \max\{-x_i, 0\}$. DEFINE

$$\lambda_j = \mu_j + s/(2n) \quad \text{FOR ALL } j \in \{-i, i\}_{i=1}^n.$$

THEN IT IS CLEAR THAT

$$\sum_{i=1}^n (\lambda_i + \lambda_{-i}) = 1,$$

AND

$$\lambda_i e_i + \lambda_{-i} e_{-i} = x_i \quad , \forall i \in \{1, \ldots, n\}.$$

1.7 ☕☕ (Random graphs with random number of vertices) Suppose that in Example 1.4.2, the number n of freshmen who arrive on campus is a random variable that has Poisson distribution with mean λ. As before, each pair of students becomes friends with probability p independent of all other pairs. Show that if $p \geq 2\ln(\lambda)/\lambda$ then there are no friendless students with probability at least $1 - 1/\lambda$.

FOR $N \sim \mathcal{P}(\lambda)$, LET B BE THE BAD EVENT THAT THERE IS AT LEAST A FRIENDLESS STUDENT. THEN, CALLING E_i THE EVENT THAT THE i-TH STUDENT IS FRIENDLESS, $P(B \mid N = n) = P(E_1 \cup \cdots \cup E_n)$. THEREFORE,

$$P(B) = \sum_{n=0}^{\infty} P(N = n) P\left(\bigcup_{i=1}^n E_i\right) \qquad \text{BY THE LAW OF TOTAL PROB.}$$

$$\leq \sum_{n=0}^{\infty} \left(e^{-\lambda} \frac{\lambda^n}{n!} \sum_{i=1}^n P(E_i) \right) \qquad \text{BY THE UNION BOUND}$$

$$= \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} n (1-p)^{n-1}$$

$$= \sum_{n=1}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} n (1-p)^{n-1}$$

$$= \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{[\lambda(1-p)]^{n-1}}{(n-1)!}$$

$$= \lambda e^{-\lambda} e^{\lambda(1-p)}$$

BY THE TAYLOR SERIES OF THE EXPONENTIAL FUNCTION

$$= \lambda e^{-\lambda p}$$

$$\leq \lambda e^{-\cancel{\lambda}(2/\cancel{\lambda})\ln \lambda}$$

$$= \lambda^{-1},$$

WHICH PROVES THE RESULT.

1.8 ☕☕ (Independent sets in random graphs) Call a group of people *independent* if no two members are friends. Suppose $n \geq 7$ students enroll in a class on high-dimensional probability, with each pair becoming friends independently with probability $1/2$. Show that, with probability at least $1 - 1/n$, this class has no independent subsets of more than $2 \log_2 n$ students.

LET $\alpha(G)$ BE THE INDEP. NUMBER OF $G(n,p)$ (i.e., THE SIZE OF THE LARGEST INDEP. SUBSET OF $G(n,p)$.) DEFINE $m := \lfloor 2\log_2 n \rfloor + 1$

$$P(\alpha(G) \geq m) \leq \binom{n}{m} 2^{-\binom{m}{2}} = \binom{n}{m} 2^{-m(m-1)/2},$$

WHERE THE INEQ. WAS OBTAINED BY APPLYING THE UNION BOUND, NOW BY THE BOUND $\binom{n}{m} \leq \left(\frac{en}{m}\right)^m$ IN EXERCISE 0.6,

$$P(\alpha(G) \geq m) \leq \left(\frac{en}{m}\right)^m 2^{-m(m-1)/2}.$$

THEN,

$$\log_2 P(\alpha(G) \geq m) \leq m[\log_2 e + \log_2 n - \log_2 m - (m-1)/2]$$

$$\le m[\log_2 e - \log_2(\lfloor 2\log_2 n\rfloor + 1) - \lfloor 2\log_2 n\rfloor/2 + \log_2 n]$$

$$= m[\log_2 e - \log_2(\lfloor \log_2 n^2\rfloor + 1) - \lfloor \log_2 n^2\rfloor/2 + \log_2 n]$$

$$\le m[\log_2 e - \log_2(\lfloor \log_2 49\rfloor + 1) - \lfloor \log_2 49\rfloor/2 + \log_2 7]$$

$$\le m[1.5 - 2.5 - 2.5 + 2.9]$$

$$= m(-0.6),$$

WHERE THE THIRD INEQUALITY IS OBTAINED BY THE FACT THAT $n \ge 7$. THEN,

$$\begin{aligned}
P(\alpha(G) \ge m) &\le 2^{-0.6m} \\
&= 2^{-0.6(\lfloor 2\log_2 n\rfloor + 1)} \\
&\le 2^{-0.6(2\log_2 n + 1)} \\
&\le 2^{-1.2\log_2 n - 0.6} \\
&= n^{-1.2}\, 2^{-0.6} \\
&\le n^{-1}.
\end{aligned}$$

1.9 ☕ (Dense random graphs have no isolated vertices) Let us refine the result of Example 1.4.2. Suppose n freshmen arrive on campus, with each pair becoming friends independently with probability p_n. Fix any $\varepsilon > 0$ and assume that

$$p_n > \frac{(1+\varepsilon)\ln n}{n} \quad \text{for every } n \in \mathbb{N}.$$

Prove that there are no friendless students with probability that converges to 1 as $n \to \infty$.

USING THE SAME NOTATION AND REASONING AS IN EXAMPLE 1.4.2 AND EXERCISE 1.7,

$$P(B) \le \sum_{i=1}^{n} P(E_i) \le n(1-p_n)^{n-1} \le n\left(1 - \frac{(1+\varepsilon)\ln n}{n}\right)^{n-1}$$

$$\le \frac{n\left(1 + \frac{\ln n^{-(1+\varepsilon)}}{n}\right)^n}{1 - \frac{(1+\varepsilon)\ln n}{n}}$$

AND TAKING LIMITS IN THE RHS,

$$\lim_{n\to\infty} \frac{n\left(1 + \frac{\ln n^{-(1+\varepsilon)}}{n}\right)^n}{1 - \frac{(1+\varepsilon)\ln n}{n}} = \lim_{n\to\infty} n n^{-1-\varepsilon} = 0.$$

1.10 ☕☕☕☕ (Sparse random graphs have isolated vertices) Let us prove a converse to Exercise 1.9. Fix any $\varepsilon > 0$ and assume that

$$p_n < \frac{(1-\varepsilon)\ln n}{n} \quad \text{for every } n \in \mathbb{N}.$$

Then there exists at least one friendless student with probability that converges to 1 as $n \to \infty$. You will prove this result using the so-called *second moment method*:

(a) Denote the number of friendless students by S_n and express it as $S_n = X_1 + \ldots + X_n$ where X_i is the indicator of the event that student i is friendless. Show that

$$\mu_n = \mathbb{E}\, S_n \to \infty.$$

Thus the expected number of friendless students is large. But this does not automatically imply that there exists even one friendless student with high probability! (Why?)

(b) Compute the second moment $\mathbb{E}\, S_n^2$ by expanding the square. Conclude that

$$\frac{\mathrm{Var}(S_n)}{\mu_n^2} \to 0.$$

(c) Use Chebyshev inequality to complete the proof.

(a) $$\lim_{n\to\infty} \mathbb{E} S_n = \lim_{n\to\infty} \mathbb{E}(X_1 + \cdots + X_n) = \lim_{n\to\infty} n(1-p_n)^{n-1} \geq \lim_{n\to\infty} n^{\varepsilon} = \infty.$$

WHERE THE FORMER TO LAST EQUALITY IS OBTAINED BY REASONING AS IN THE PREVIOUS EXERCISE. TO SEE THAT THE PROBABILITY OF EVEN ONE FRIENDLESS STUDENT IS SMALL, OBSERVE THAT CORRELATION BETWEEN THE X_i'S CAN AFFECT THIS PROB. FOR INSTANCE, IF $X_1 \sim \mathrm{Ber}(p_n)$ AND $X_i = X_1$ a.s. FOR ALL $i = 2, 3, \ldots, n$, THEN

$$\mathbb{E} S_n = n p_n = (1-\varepsilon)\ln n \xrightarrow{n\to\infty} \infty$$

BUT

$$P(S_n = n) = p_n \xrightarrow{n\to\infty} 0.$$

(b)

$$\begin{aligned} \mathbb{E} S_n^2 &= \mathbb{E}[(X_1 + \cdots + X_n)(X_1 + \cdots + X_n)] \\ &= \mathbb{E} \sum_{i=1}^n X_i^2 + \sum_{1 \le i \ne j \le n} \mathbb{E} X_i X_j \\ &= n(1-p_n)^{n-1} + \sum_{1 \le i \ne j \le n} P[i,j \text{ ARE FRIENDLESS}] \end{aligned}$$

$$= n(1-p_n)^{n-1} + \sum_{1\le i\ne j\le n} P(i \text{ IS FRIENDLESS}) P(j \text{ IS FRIENDLESS} \mid i \text{ IS FRIENDLESS})$$

$$= n(1-p_n)^{n-1} + \sum_{i\ne j} (1-p_n)^{n-1} (1-p_{n-1})^{n-2}$$

SINCE j CANNOT BE FRIENDS WITH i. ALSO FROM THE SECOND TERM IN THE LAST EQUALITY, WE SEE THAT X_i AND X_j ARE ALMOST UNCORRELATED (IN FACT, THEY ARE ASYMPTOTICALY UNCORRELATED). THUS,

$$\mathbb{E} S_n^2 = n(1-p_n)^{n-1} + \sum_{i\ne j} (1-p_n)^{n-1} (1-p_{n-1})^{n-2}$$

$$= n(1-p_n)^{n-1} + 2\binom{n}{2} (1-p_n)^{n-1} (1-p_{n-1})^{n-2}$$

$$= \underbrace{n(1-p_n)^{n-1}}_{\mathbb{E} S_n} + \underbrace{n(1-p_n)^{n-1}}_{\mathbb{E} S_n} \underbrace{(n-1)(1-p_{n-1})^{n-2}}_{\mathbb{E} S_{n-1}} .$$

THEN,

$$V(S_n) = \mathbb{E} S_n + \mathbb{E} S_n \, \mathbb{E} S_{n-1} - (\mathbb{E} S_n)^2,$$

AND THE LAST TWO TERMS CANCEL EACH OTHER ASYMPTOTICALY. THEREFORE,

$$\lim_{n\to\infty} \frac{V(S_n)}{(\mathbb{E} S_n)^2} = \lim_{n\to\infty} \frac{1}{\mathbb{E} S_n} = 0.$$

(c)

$$P(S_n = 0) \le P(S_n \le 0)$$

$$= P(S_n - \mu_n \le -\mu_n)$$

$$\le P(|S_n - \mu_n| \ge \mu_n)$$

$$\le V(S_n)/\mu_n^2 \xrightarrow{n\to\infty} 0$$

1.11 ☕☕ (Monotonicity of the L^p norm)

(a) Let X be a random variable. Show that $\|X\|_{L^p}$ is an increasing function in p:

$$\|X\|_{L^p} \le \|X\|_{L^q} \quad \text{for any } 0 \le p \le q \le \infty.$$

(b) Demonstrate that the inequality in part (a) can not be reversed: for any $0 \le p < q \le \infty$, find an example of a random variable X with $\|X\|_{L^p} < \infty$ and $\|X\|_{L^q} = \infty$.

(a) LET $f(x) = x^{q/p}$. BY JENSEN'S INEQ.,

$$\|X\|_{L^p}^q = (\mathbb{E}|X|^p)^{q/p} = f(\mathbb{E}|X|^p) \leq \mathbb{E} f(|X|^p) = \mathbb{E}|X|^q.$$

TAKING THE q-TH ROOT ON BOTH SIDES,

$$\|X\|_{L^p} \leq \|X\|_{L^q}.$$

(b) LET $m \in (p,q)$ AND CONSIDER A R.V. X FULLY SUPPORTED ON $\mathbb{R}^+$ WITH DENSITY $f(x) \sim x^{-m}$. THEN

$$\mathbb{E} X^p = \int_{\mathbb{R}^+} x^p f(x) dx \sim \int_{\mathbb{R}^+} (x^p/x^m) \, dx < \infty$$

AND

$$\mathbb{E} X^q = \int_{\mathbb{R}^+} x^q f(x) dx = \int_{\mathbb{R}^+} (x^q/x^m) dx \doteq \infty.$$

AS FOR $q=\infty$, SINCE X IS NOT BOUNDED, $\|X\|_{L^\infty} = \infty$.

1.12 ☕ (Interpolation between L^1 and L^∞) We know the L^p norm of any random variable X is bounded by the L^∞ norm. We can get an even better bound if we also know that the L^1 norm of X is small. Show that

$$\|X\|_{L^p} \leq \|X\|_{L^1}^{\frac{1}{p}} \|X\|_{L^\infty}^{1-\frac{1}{p}} \quad \text{for any } 1 < p < \infty.$$

CALL $M := \|X\|_{L^\infty}$. THEN

$$\mathbb{E}|X|^p = \mathbb{E}(|X||X|^{p-1}) \leq \mathbb{E}(|X| M^{p-1}) = \mathbb{E}|X| M^{p-1} = \|X\|_{L^1} \|X\|_{L^\infty}^{p-1}.$$

AND TAKING THE p-TH ROOT, THE RESULT FOLLOWS.

1.13 ☕☕☕ (Expectation of a maximum) Let $X_1, \ldots, X_n$ be nonnegative random variables.

(a) Prove that

$$\max_{i \leq n} \mathbb{E} X_i \leq \mathbb{E} \max_{i \leq n} X_i \leq n \cdot \max_{i \leq n} \mathbb{E} X_i.$$

(b) Demonstrate that both inequalities in part (a) may be optimal. Specifically, find random variables $X_1, \ldots, X_n$ satisfying $\max_i \mathbb{E} X_i = \mathbb{E} \max_i X_i > 0$ and random variables Y_i satisfying $\mathbb{E} \max_i Y_i = n \cdot \max_i \mathbb{E} Y_i > 0$.

(c) Demonstrate that the upper bound in part (a) may be approximately optimal even for independent random variables. Specifically, find independent random variables $X_1, \ldots, X_n$ satisfying $\mathbb{E} \max_i X_i > cn \cdot \max_i \mathbb{E} X_i$, where $c > 0$ is an absolute constant.[5]

(a) THE FIRST INEQ. IS A CONSEQUENCE OF EXERCISE 1.2 AND JENSEN'S INEQ.

AS FOR THE SECOND, ASSUME THAT $\mathbb{E}\max X_i > n \max \mathbb{E}X_i$. THEN

$$\mathbb{E}\sum_{i=1}^{n} X_i \geq \mathbb{E}\max_{i\leq n} X_i > n \max_{i \leq n} \mathbb{E}X_i \geq \sum_{i=1}^{n} \mathbb{E}X_i = \mathbb{E}\sum_{i=1}^{n} X_i,$$

A CONTRADICTION BECAUSE AN INNER INEQ. IS STRICT.

(b) LET $X_1 = \cdots = X_n = C \in \mathbb{R}^+$ BE DEGENERATE R.V.'S. THEN $\max_{i\leq n} \mathbb{E}X_i = \mathbb{E}\max_{i \leq n} X_i$.

LET U BE A UNIFORM R.V. OVER $I = \{1, \ldots, n\}$ AND DEFINE $Y_i = n\mathbb{1}_{\{i\}}$, $i \in I$.

SUPPOSE U IS REALIZED AT $u_0 \in I$. THEN $Y_{u_0} = n$ AND $Y_i = 0$ FOR ALL $i \neq u_0$.

THEREFORE, $\max_{i\in I} Y_i = Y_{u_0} = n$, AND

$$\mathbb{E}\max_{i \in I} Y_i = \mathbb{E}Y_{u_0} = n.$$

ON THE OTHER HAND, FOR ALL $i \in I$, $\mathbb{E}Y_i = 1$. THEREFORE

$$\mathbb{E}\max_{i\in I} Y_i = n \cdot \max_{i \in I} \mathbb{E}Y_i,$$

AS REQUIRED.

(c) LET $X_1, \ldots, X_n$ BE INDEP. BERNOULLI TRIALS WITH PARAMETER $p = p_n$. WHEN $n=1$, THE RESULT IS TRIVIAL. SO LET'S ASSUME $n \geq 2$. ACCORDINGLY, $\mathbb{E}X_i = p_n$ FOR $i = 1, \ldots, n$, AND $\max_{i \leq n} \mathbb{E}X_i = p_n$. ON THE OTHER HAND, LET $Y = \max_{i\leq n} X_i$. THEN

$$\begin{aligned}
\mathbb{E}\max_{i \leq n} X_i &= 1 \cdot P(X_i = 1 \text{ FOR AT LEAST ONE } i) + 0 \cdot P(\cancel{X_i = 0} \text{ FOR ALL } i) \\
&= 1 \cdot (1 - P(X_i = 0 \text{ FOR ALL } i)) \\
&= 1 - (1 - p_n)^n.
\end{aligned}$$

THUS, THE DESIRED INEQ. IS SATISFIED IF

$$\mathbb{E}\max_{i\le n} X_i = 1-(1-p_n)^n \ge c\cdot n p_n = cn \max_{i\le n}\mathbb{E}X_i, \qquad (*)$$

FOR SOME $c>0$. NOW, OBSERVE THAT

$$\begin{aligned} 1-(1-p_n)^n &\ge 1-e^{-np_n} \\ &\ge 1-\left(1-np_n+(np_n)^2/2\right) \\ &= np_n-(np_n)^2/2 \end{aligned}$$

THEREFORE, $(*)$ IS SATISFIED IF THERE IS $c>0$ S.T

$$\begin{aligned} np_n - n^2p_n^2/2 \ge cnp_n &\iff -n^2p_n^2/2 \ge (c-1)np_n \\ &\iff -np_n/2 \ge c-1 \\ &\iff -np_n \ge 2c-2 \\ &\iff np_n \le 2(c-1). \end{aligned}$$

THUS, IF $p_n = n^{-1}$, $c \ge 3/2$ SATISFIES $(*)$.

1.14 ☕☕ Let $X_1,\ldots,X_n$ be nonnegative random variables. Prove that for any $1\le p<\infty$, we have

$$\left(\sum_{i=1}^n (\mathbb{E}\,X_i)^p\right)^{1/p} \le \mathbb{E}\left(\sum_{i=1}^n X_i^p\right)^{1/p} \le \left(\sum_{i=1}^n \mathbb{E}\,(X_i^p)\right)^{1/p}.$$

LET $g(x) = x^{1/p}$, WHICH IS CONCAVE. BY JENSEN'S INEQ.,

$$\mathbb{E}\left(\sum_{i=1}^n X_i^p\right)^{1/p} = \mathbb{E}\, g\left(\sum_{i=1}^n X_i^p\right) \le g\left(\mathbb{E}\sum_{i=1}^n X_i^p\right) = \left(\mathbb{E}\sum_{i=1}^n X_i^p\right)^{1/p},$$

WHICH ESTABLISHES THE SECOND INEQ. FOR THE FIRST INEQ., OBSERVE THAT

$$\left(\sum_{i=1}^n (\mathbb{E}X_i)^p\right)^{1/p} = \|\mathbb{E}X\|_p \le \mathbb{E}\|X\|_p = \mathbb{E}\left(\sum_{i=1}^n X_i^p\right)^{1/p}.$$

1.15 ☕ (Integrated tail formulas) Prove the following more general versions of Lemma 1.6.1.

(a) Let X be any random variable, not necessarily nonnegative. Then

$$\mathbb{E}\, X = \int_0^\infty \mathbb{P}\{X > t\}\, dt - \int_{-\infty}^0 \mathbb{P}\{X < t\}\, dt.$$

(b) Let X be a nonnegative random variable. Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be an increasing, differentiable function satisfying $f(0) = 0$. Then

$$\mathbb{E}\, f(X) = \int_0^\infty \mathbb{P}\{X > t\} f'(t) dt.$$

(c) Let X be a random variable, not necessarily nonnegative. Deduce that for every $p \in (0, \infty)$ we have

$$\mathbb{E}|X|^p = \int_0^\infty \mathbb{P}\{|X| > t\} p t^{p-1}\, dt$$

(a) LET $X = X^+ - X^-$, WHERE $X^+ = \max\{0, X\}$ AND $X^- = \max\{0, -X\}$. SINCE BOTH X^+ AND X^- ARE NONNEGATIVE, THE RESULT FOLLOWS.

(b) LET $g_X(x)$ BE THE DENSITY OF X. THEN

$$\int_{\mathbb{R}^+} P[X > t] f'(t) dt = \int_0^\infty \int_t^\infty g(x) f'(t) dx\, dt$$

$$= \int_0^\infty \int_0^x g(x) f'(t) dt\, dx$$

$$= \int_0^\infty g(x) \int_0^x f'(t) dt\, dx$$

$$= \int_0^\infty g(x) f(x) dx$$

$$= \mathbb{E} f(X)$$

(c) DIRECT FROM (b).

1.16 ☕☕ (Paley-Zygmund inequality) Markov inequality says a random variable is unlikely to be much bigger than its expectation. But what about the reverse? Can a nonnegative random variable be much smaller than its expectation with high probability? In general, yes (example?), but not if the second moment isn't too large. Let X be a nonnegative random variable with finite variance. Show that for any $\varepsilon \in [0, 1]$:

$$\mathbb{P}\{X > \varepsilon \mathbb{E} X\} \geq (1-\varepsilon)^2 \frac{(\mathbb{E} X)^2}{\mathbb{E}[X^2]}.$$

LET $E = \{X > \varepsilon \mathbb{E} X\}$. THEN

$$\begin{aligned} \mathbb{E} X &= \mathbb{E}[X \mathbb{1}_E] + \mathbb{E}[X \mathbb{1}_{E^c}] \\ &\leq \mathbb{E}[X \mathbb{1}_E] + \varepsilon \mathbb{E} X \\ &\leq (\mathbb{E} X^2)^{1/2} (\mathbb{E} \mathbb{1}_E^2)^{1/2} + \varepsilon \mathbb{E} X \\ &= (\mathbb{E} X^2)^{1/2} \, P(X > \varepsilon \mathbb{E} X)^{1/2} + \varepsilon \mathbb{E} X, \end{aligned}$$

AND THE RESULT IS OBTAINED.

1.17 ☕☕☕ (Comparison of the ℓ^p norms) Let $0 \leq p \leq q \leq \infty$.

(a) Prove that for any vector $x \in \mathbb{R}^n$ we have

$$\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{q}} \|x\|_q.$$

(b) Demonstrate that both inequalities in part (a) can be optimal. Specifically, find nonzero vectors $x, y \in \mathbb{R}^n$ satisfying $\|x\|_p = \|x\|_q$ and $\|y\|_p = n^{\frac{1}{p} - \frac{1}{q}} \|y\|_q$.

(a) LET $x = (x_1, \ldots, x_n)$. BY HOMOGENEITY (THAT $\|\lambda x\| = |\lambda| \|x\|$ FOR ANY NORM $\|\cdot\|$), IT IS ENOUGH TO PROVE THE RESULT FOR $y = x / \|x\|_p$. THEN

$$\begin{aligned} y = x/\|x\|_p \quad &\Rightarrow \|y\|_p = 1 \\ &\Rightarrow \sum_{i=1}^n |y_i|^p = 1 \\ &\Rightarrow |y_i| \leq 1, \quad 1 \leq i \leq n \\ &\Rightarrow |y_i|^q \leq |y_i|^p, \quad 1 \leq i \leq n \\ &\Rightarrow \sum_{i=1}^n |y_i|^q \leq \sum_{i=1} |y_i|^p = 1 \\ &\Rightarrow \|y\|_q \leq 1 = \|y\|_p. \end{aligned}$$

AS FOR $q=\infty$, $|x_i|^p \le \sum_{i=1}^{n}|x_i|^p = \|x\|_p^p$ FOR ALL i. THEREFORE

$$\|x\|_\infty = \max|x_i| \le \|x\|_p.$$

(b) LET $x=(1,0,\dots,0)$. THEN $\|x\|_p = \|x\|_q$.

LET $y=(1,\dots,1)$. THEN $\|y\|_p = n^{1/p}$ AND $\|y\|_q = n^{1/q}$

$$\begin{aligned} &= n^{1/p}\cdot 1 \\ &= n^{1/p}n^{-1/q}n^{1/q} \\ &= n^{\frac{1}{p}-\frac{1}{q}}\|y\|_q \end{aligned}$$

1.18 ☕ (The ℓ^∞ norm is the limit of the ℓ^p norms) Consider any vector $x \in \mathbb{R}^n$.

(a) Prove that

$$\|x\|_p \to \|x\|_\infty \text{ as } p \to \infty.$$

(b) In fact, p does not need to be too large for the ℓ^p norm to get reasonably close to the ℓ^∞ norm. Show that if $p \ge \ln n$, then

$$\|x\|_\infty \le \|x\|_p \le e\|x\|_\infty.$$

(a) LET $M = \max|x_i| = \|x\|_\infty$, THEN

$$\|x\|_\infty \le \|x\|_p = \left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p} \le n^{1/p}M \xrightarrow{p\to\infty} M = \|x\|_\infty, \qquad (*)$$

WHERE THE FIRST INEQ. WAS OBTAINED IN EXERCISE 1.17(a).

(b) $p \ge \ln n \Leftrightarrow e^p \ge n$. THEREFORE, $(*)$ BECOMES

$$\|x\|_\infty \le \|x\|_p \le n^{1/p}M \le eM = e\|x\|_\infty.$$

1.19 ☕☕☕ (Duality of the ℓ_p norms) Let $p, p' \in [1,\infty]$ be conjugate exponents.

(a) Show that Hölder inequality is tight: for any vector x there exists a vector $y \neq 0$ for which $\langle x,y\rangle = \|x\|_p\|y\|_{p'}$.

(b) Conclude that for every vector $x \in \mathbb{R}^n$, we have

$$\max\left\{\langle x,y\rangle : \ y \in B^n_{p'}\right\} = \|x\|_p.$$

(a)

- $p=1$, $p'=\infty$

- ASSUME THAT $x_i \geq 0$ FOR ALL $i \in \{1,\dots,n\}$. THEN, LETTING $y = \mathbb{1} = (1,\dots,1)$,

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i = \|x\|_1 \cdot \|y\|_\infty .$$

- MORE GENERALLY, IF $x_i < 0$, TAKE $y_i = -1$
  $x_i \geq 0$, TAKE $y_i = 1$.

THEN $\langle x,y\rangle = \sum_{i=1}^{n} |x_i| = \|x\|_1 \cdot \|y\|_\infty .$

- $p=\infty$, $p'=1$

IF $x_i \geq 0$ FOR ALL $i \in \{1,\dots,n\}$, TAKE $\|x\|_\infty = x_k$ AND MAKE $y = (0,\dots,0,1,0,\dots,0)$ (WITH A 1 IN THE k-TH COORDINATE AND 0 ELSEWHERE. THEN

$$\langle x,y\rangle = x_k = \|x\|_\infty = \|x\|_\infty \|y\|_1 .$$

IF $\|x\|_\infty = |x_k|$ AND $x_k < 0$,

$$\langle x,-y\rangle = |x_k| = \|x\|_\infty = \|x\|_\infty \|-y\|_1 .$$

- $p, p' \in (1,\infty)$

IF $x_i \geq 0$ FOR ALL $x_i \in x$,

MAKE $y_i = x_i^{\,p/p'}$. THEN

$$\|y\|_{p'} = \left(\sum_{i=1}^{n} y_i^{p'}\right)^{1/p'} = \left(\sum_{i=1}^{n} x_i^{p}\right)^{1/p'} = \|x\|_p^{p/p'} = \|x\|_p^{p-1}$$

AND

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i x_i^{p/p'} = \sum_{i=1}^{n} x_i^{p} = \|x\|_p^{p} = \|x\|_p \|x\|_p^{p-1} = \|x\|_p \|y\|_{p'}$$

IF $x_i < 0$, TAKE $y_i = -|x_i|^{p/p'}$.

(b) THIS IS A CONSEQUENCE OF HÖLDER'S INEQ AND THE DEF OF THE COSINE OF THE ANGLE BETWEEN x AND y:

$$\max\{\langle x,y\rangle : y\in B_{p'}^n\} = \max\{\langle x,y\rangle : \|y\|_{p'}\le 1\}$$
$$\le \max\{|\langle x,y\rangle| : \|y\|_{p'}\le 1\}$$
$$\le \max\{\|x\|_p\|y\|_{p'} : \|y\|_{p'}=1\}$$
$$= \|x\|_p$$

THUS, $\max\{\langle x,y\rangle : y\in B_{p'}^n\}\le\|x\|_p$. TO ATTAIN THE UPPER BOUND, MAKE $y^* = x/\|x\|_{p'}$. THEN $\|y^*\|_{p'}=1$ AND $y^* = \arg\max_{y:\|y\|_{p'}=1}\langle x,y\rangle$, BY DEF OF THE COSINE.