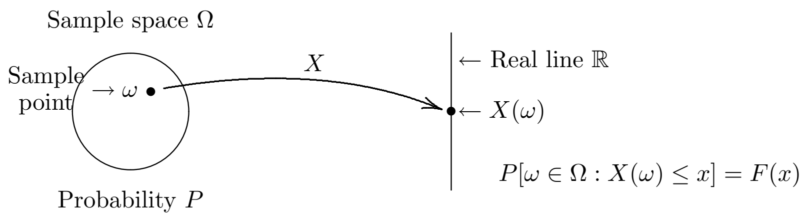
TOPIC. Expectations. This section deals with the notion of the expected value of a random variable. We start with some definitions and examples, then give some ways of thinking about expected values, and then present some properties of expectation along with examples.

Definitions. Let Ω be a sample space, P a probability measure on Ω, X a real-valued random variable on Ω with distribution function F. This situation is illustrated below:



We are going to define $E(X)$, the ***expected value*** of $X(\omega)$ when the sample point ω is chosen at random from Ω according to P. An alternative notation for $E(X)$ is $\int_{\omega \in \Omega} X(\omega)\, P(d\omega)$, or simply $\int X\, dP$.

Consider first the case where X is nonnegative: $X(\omega) \geq 0$ for all $\omega \in \Omega$. If X is discrete, taking finitely or countably many values $x_1, x_2, \ldots$ with corresponding probabilities $f(x_1), f(x_2), \ldots$ (here f denotes the probability mass function of X), one takes

$$E(X) := \sum\nolimits_k x_k f(x_k). \tag{1}$$

If X is continuous with density f, one takes

$$E(X) := \int_0^\infty x\, f(x)\, dx. \tag{2}$$

Formulas (1) and (2) are each special cases of the general definition

$$E(X) := \int_0^\infty x\, dF(x) \ \Big(:= \lim_{n\to\infty} \int_0^n x\, dF(x)\Big) \tag{3}$$

where the integral is taken to be a Riemann-Stieltjes integral. For this course you don't need to know much about Riemann-Stieltjes integration; you can just think of the RHS of (3) as a generic way of writing the RHSs of (1) and (2). We don't require the sum and integrals in (1)–(3) to converge to a finite value; $E(X) = \infty$ is allowed, and happens (see Example 1 (b) below).

Example 1. (a) Suppose Z is a standard normal random variable, and consider

$$Z^+ := \max(Z, 0) = \begin{cases} Z, & \text{if } Z \geq 0, \\ 0, & \text{if } Z < 0. \end{cases}$$

Z^+ is a nonnegative random variable. Its distribution has a lump of mass of size 1/2 at 0 and density $\phi(z) = e^{-z^2/2}/\sqrt{2\pi}$ over the interval $(0, \infty)$. Hence

$$\begin{aligned} E(Z^+) &= 0 \times P[Z^+ = 0] + \int_0^\infty z\phi(z)\, dz \\ &= 0 \times \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^\infty z e^{-z^2/2}\, dz = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-y}\, dy = \frac{1}{\sqrt{2\pi}}. \end{aligned} \tag{4}$$

In particular $E(Z^+)$ is finite.

(b) Suppose C is a standard Cauchy random variable, with density $f(x) = 1/(\pi(1+x^2))$ on $\mathbb{R}$. Then $C^+ = \max(C, 0)$ is a nonnegative random variable with expectation

$$\begin{aligned} E(C^+) &= 0 \times \frac{1}{2} + \int_0^\infty \frac{x}{\pi(1+x^2)}\, dx \\ &= \frac{1}{2\pi} \int_0^\infty \frac{1}{1+y}\, dy = \frac{1}{2\pi} \log(1+y)\Big|_0^\infty = \infty. \end{aligned} \tag{5}$$

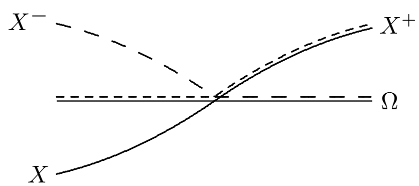
Note that $E(C^+)$ is infinite. •

Now consider the case where X can take both positive and negative values. Define random variables X^+ and X^- on Ω by setting

$$X^+(\omega) = \max(X(\omega), 0) \quad = \begin{cases} X(\omega), & \text{if } X(\omega) \geq 0, \\ 0, & \text{otherwise}, \end{cases} \tag{6_+}$$

$$X^-(\omega) = \max(-X(\omega), 0) = \begin{cases} -X(\omega), & \text{if } X(\omega) \leq 0, \\ 0, & \text{otherwise}, \end{cases} \tag{6_-}$$

for each $\omega \in \Omega$, as illustrated below:



X^+ is called the ***positive part*** of X, and X^- the ***negative part***. Note that X^+ and X^- are nonnegative random variables and that

$$X(\omega) = X^+(\omega) - X^-(\omega) \quad \text{and} \quad |X(\omega)| = X^+(\omega) + X^-(\omega)$$

for all $\omega \in \Omega$; these identities are written more concisely as $X = X^+ - X^-$ and $|X| = X^+ + X^-$.

One says that ***X has an expectation***, or that ***E(X) exists***, or that ***X is quasi-integrable*** if at least one of $E(X^+)$ and $E(X^-)$ is finite; in that case the ***expected value***, or ***mean***, ***of X*** is taken to be

$$E(X) := E(X^+) - E(X^-) \tag{7}$$

with the convention that $\infty - x = \infty$ and $x - \infty = -\infty$ for any nonnegative real number x. One says that ***X is integrable***, or that ***X has a finite expectation***, if both $E(X^+)$ and $E(X^-)$ are finite, or, equivalently, if $E(|X|)$ is finite. There are random variables X for which $E(X^+) = \infty = E(X^-)$; for such X's, $E(X)$ is not defined.

Example 2. (a) Suppose Z is a standard normal random variable. By Example 1 (a), we have $E(Z^+) = c := 1/\sqrt{2\pi} < \infty$. Since Z^- and Z^+ have the same distribution, we also have $E(Z^-) = c$. Since both $E(Z^+)$ and $E(Z^-)$ are finite, Z is integrable; its (finite) expectation is

$$E(Z) = E(Z^+) - E(Z^-) = c - c = 0.$$

(b) Suppose C is a standard Cauchy random variable. By Example 1 (b) and symmetry, we have $E(C^+) = \infty = E(C^-)$. Thus C does not have an expectation, finite or otherwise.

(c) As in (b), suppose C is standard Cauchy. Put $X = C^+$. Then

$$X^+ = X = C^+ \Longrightarrow E(X^+) = \infty \quad \text{and}$$
$$X^- = 0 \Longrightarrow E(X^-) = 0.$$

Consequently X has an expectation, namely

$$E(X) = E(X^+) - E(X^-) = \infty - 0 = \infty.$$

In this case $E(X)$ exists, but is infinite; this is an example of a random variable that is quasi-integrable, but not integrable.

(d) Suppose X is a continuous random variable with density f on $\mathbb{R}$ such that the integral $\int_{-\infty}^{\infty} x f(x)\,dx$ is absolutely convergent. Since

$$\begin{aligned} &E(X^+) + E(X^-) \\ &\quad = \int_0^\infty x f(x)\,dx + \int_{-\infty}^0 (-x) f(x)\,dx = \int_{-\infty}^\infty |x| f(x)\,dx < \infty, \end{aligned}$$

X is integrable with finite expectation

$$\begin{aligned} &E(X) = E(X^+) - E(X^-) \\ &\quad = \int_0^\infty x f(x)\,dx - \int_{-\infty}^0 (-x) f(x)\,dx = \int_{-\infty}^\infty x f(x)\,dx. \end{aligned} \tag{8}$$

When it applies, (8) can be used to compute $E(X)$ directly, without first computing $E(X^+)$ and $E(X^-)$. •

The strong law of large numbers (SLLN). Why is $E(X)$ important? One of the main reasons is:

Theorem 1 (The SLLN). *Suppose X_1, X_2, $\ldots$ is an infinite sequence of independent random variables, each distributed like a random variable X. Put*

$$S_n = X_1 + X_2 + \cdots + X_n$$

for each $n \in \mathbb{N}$. If X has an expectation $E(X) = \mu$ (possibly $\pm\infty$) then

$$P[S_n/n \text{ converges to } \mu \text{ as } n \to \infty] = 1. \tag{9_1}$$

On the other hand, if X does not have an expectation, then

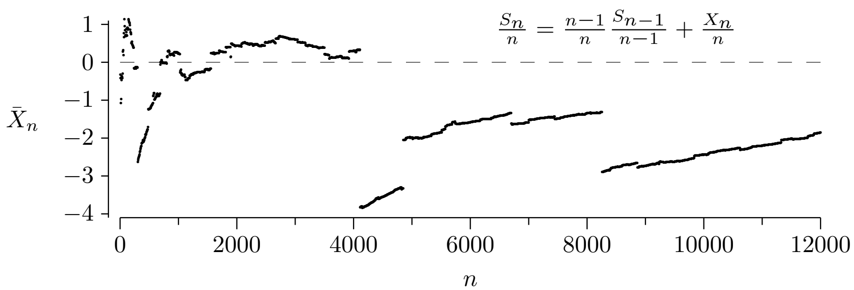
$$P[\limsup_n |S_n/n| = \infty] = 1; \tag{9_2}$$

if in addition X is symmetric, then

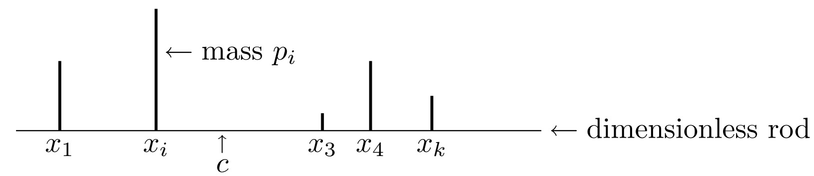
$$P[\liminf_n S_n/n = -\infty] = 1 = P[\limsup_n S_n/n = \infty]. \tag{9_3}$$

In other words, if the "population mean" μ exists, then the sample means $\bar{X}_n = S_n/n$ will converge to it almost surely as the sample size n tends to infinity; but if μ does not exist, the sample means $\bar{X}_n$ will behave very badly as $n \to \infty$, as illustrated in Figure 1 below. The proof of the SLLN is not easy; we won't go into it here (but see Exercises 10 and 11 for some special cases).

Figure 1: A graph of S_n/n versus n for a random sample of size 12000 from the standard Cauchy distribution.



***E*(*X*) as a measure of location.** The expected value of X is often used as a measure of the location of the distribution of X. To understand why, consider the case where X takes finitely many values x_1, x_2, $\ldots$, x_k with corresponding probabilities p_1, p_2, $\ldots$, p_k. We can represent the distribution of X by a physical system in which a masses of weight p_i are placed above the points x_i for $i = 1, \ldots, k$ on a dimensionless rod, as illustrated below:



Consider the ***center of gravity*** of this mass system, i.e., the point c at which the rod would balance if it were pivoted there. According to physics, c must satisfy the so-called ***balancing equation***

$$\sum_{i=1}^{k} p_i(x_i - c) = 0.$$

Since

$$\sum_{i=1}^{k} p_i x_i = E(X) \quad \text{and} \quad \sum_{i=1}^{k} p_i = 1$$

the solution to the balancing equation is

$$c = \frac{\sum_{i=1}^{k} p_i x_i}{\sum_{i=1}^{k} p_i} = E(X). \tag{10}$$

In general, for any integrable random variable X, the center of gravity of the distribution of X is $c = E(X)$. There is an important corollary: moving a little bit of probability mass a long way from its initial position has a big effect on the expected value of X.

The expected value of a transformation of X. Suppose $Y = t(X)$ is a transformation t of X. The expected value of Y can be expressed directly in terms of the distribution of X. To see how, consider the case where X is continuous with density f on $(-\infty, \infty)$ and the transformation t is regular from $(-\infty, \infty)$ to $(0, \infty)$. Since Y has density

$$f_Y(y) = f_X\big(u(y)\big)|u'(y)|$$

where $u = t^{-1}$ is the inverse of t (see 3.19), Y has expectation

$$\begin{aligned}\int_{y>0} y f_Y(y)\, dy &= \int_{y>0} t\big(u(y)\big) f_X\big(u(y)\big)|u'(y)|\, dy \\ &\underset{x=u(y)}{=} \int_{-\infty<x<\infty} t(x) f_X(x)\, dx \qquad (11)\end{aligned}$$

by (3.18). The point is that you can find $E(Y)$ from (11) without having to first work out the distribution of Y. This important fact is true in general.

Theorem 2. *Let X be an arbitrary (not necessarily continuous) random variable) and let $Y = t(X)$ for an arbitrary (not necessarily regular) transformation t. Put*

$$\tau_+ = \int t^+(x)\, dF_X(x) \quad \text{and} \quad \tau_- = \int t^-(x)\, dF_X(x).$$

Then Y has an expectation if and only if at least one of τ_+ and τ_- is finite, in which case

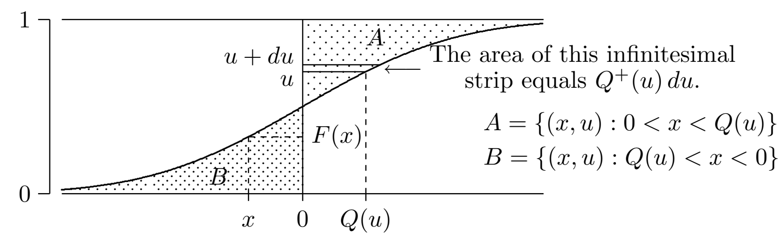
$$E(Y) = \tau_+ - \tau_- = \int t(x)\; dF_X(x). \qquad (12)$$

With an appropriate definition of the integral, this formula is valid even if X is a (multi-dimensional) random vector. These results are proved in Stat 381.

Expressing $E(X)$ in terms of Q and F. Let X be a random variable with quantile function Q and distribution function F. $E(X)$ can be expressed directly in terms of Q, and also directly in terms of F. To see how, let U be a standard uniform random variable, with density $f_U(u) = I_{(0,1)}(u)$. Since X and $Q(U)$ have the same distribution by the IPT Theorem (Theorem 1.5), so do X^+ and $Q^+(U)$, whence

$$E(X^+) = E\big(Q^+(U)\big) \underset{\text{by (12)}}{=} \int_0^1 Q^+(u)\, du.$$

The integral here is the area $|A|$ of the region $A = \{(x,u) : 0 < x < Q(u)\}$ indicated below:



By slicing A into infinitesimal vertical strips instead of horizontal ones, we can also compute its area as

$$|A| = \int_0^\infty \big(1 - F(x)\big)\, dx = \int_0^\infty P[X > x]\, dx = \int_0^\infty P[X \geq x]\, dx.$$

Similarly

$$E(X^-) = E\big(Q^-(U)\big) = \int_0^1 Q^-(u)\, du = |B| = \int_{-\infty}^0 F(x)\, dx,$$

where $B = \{(x,u) : Q(u) < x < 0\}$. This proves:

Theorem 3. *Let X be a random variable with df F and quantile function Q, and let A and B be defined as above. X is quasi-integrable if and only if at least one of $|A|$ and $|B|$ is finite, and then*

$$\begin{aligned}E(X) &= |A| - |B| \\ &= \int_0^1 Q(u)\, du = \int_0^\infty \big[-F(-x) + \big(1 - F(x)\big)\big]\, dx. \qquad (13)\end{aligned}$$

If X is quasi-integrable, then $E(X) = \int_0^\infty \big[-F(-x) + \big(1 - F(x)\big)\big]\,dx$.

Example 3. (a) For any random variable

$$E(|X|) = \int_0^\infty P[\,|X| \geq x\,]\,dx. \tag{14}$$

Consequently X is integrable if and only if this integral is convergent. For a standard Cauchy random variable C, $P[\,|C| \geq x\,] \sim 2/(\pi x)$ as $x \to \infty$, so the integral diverges; this is another way to see that C is not integrable. Note that if X is integer valued, then

$$E(|X|) = \sum\nolimits_{n=1}^{\infty} P[\,|X| \geq n\,]. \tag{15}$$

(b) Let X be a standard exponential random variable, with density $f(x) = e^{-x} I_{(0,\infty)}(x)$ on $\mathbb{R}$. Note that X takes on only nonnegative values. We have

$$F(x) = \int_0^x e^{-\xi}\,d\xi = 1 - e^{-x}$$

for $x \geq 0$, and

$$Q(u) = F^{-1}(u) = -\log(1-u)$$

for $0 < u < 1$. By calculus

$$\int_0^\infty x\,f(x)\,dx = \int_0^\infty xe^{-x}\,dx = \Gamma(2) = 1,$$
$$\int_0^\infty \big(1 - F(x)\big)\,dx = \int_0^\infty e^{-x}\,dx = \Gamma(1) = 1,$$
$$\int_0^1 Q(u)\,du = \int_0^1 -\log(1-u)\,du = \int_0^1 -\log(v)\,dv = 1.$$

Of course, all three integrals had to be the same, since they each give the value of $E(X)$. •

Properties of *E*. We state without proof some basic properties of the expectation operator E. These properties are proved (perhaps under some further integrability assumptions) in elementary texts in the discrete and continuous case; they are proved in general in Stat 381.

Theorem 4. *Expectation has the following properties.*

E_+: *If two random variables X and Y each have finite expectations, then so does $X + Y$, and*

$$E(X+Y) = E(X) + E(Y). \tag{16}$$

More generally, if $E(X)$ and $E(Y)$ exist (possibly as $\pm\infty$) and if the sum $E(X) + E(Y)$ is defined (i.e., is not of the form $+\infty - \infty$ or $-\infty + \infty$), then $E(X+Y)$ exists and is given by (16).

E_c: *If X has an expectation and c is a finite real number, then cX has an expectation, given by*

$$E(cX) = cE(X). \tag{17}$$

$\mathsf{E}_\leq$: *Suppose X and Y are two random variables such that $X \leq Y$ (i.e., $X(\omega) \leq Y(\omega)$ for all sample points ω). Then*

$$E(X) \leq E(Y) \tag{18}$$

provided both expectations exist. If in addition the expectations are equal and finite, then $P[\,X = Y\,] = 1$.

E_I: *Suppose $X_1, X_2, \ldots, X_n$ are independent random variables. Then the product $X_1 X_2 \cdots X_n$ has an expectation provided: (a) all the X_k's are nonnegative, or (b) all the X_k's are integrable. In both of these cases,*

$$E(X_1 X_2 \cdots X_n) = E(X_1) E(X_2) \cdots E(X_n). \tag{19}$$

In case (a) the product on the right-hand side is to be evaluated using the rule $\infty \times c = c \times \infty$ equals ∞ if $0 < c \leq \infty$, and equals 0 if $c = 0$.

Example 4. (a) Let $X \sim \text{Gamma}(r, \lambda)$, with density $\lambda^r x^{r-1} e^{-\lambda x}/\Gamma(r)$ for $x > 0$. Then $Y = \lambda X \sim \text{Gamma}(r, 1)$, so $E(X) = E(Y)/\lambda$. Moreover

$$E(Y) = \frac{1}{\Gamma(r)} \int_0^\infty y\, y^{r-1} e^{-y}\, dy$$
$$= \frac{\Gamma(r+1)}{\Gamma(r)} \Big(\frac{1}{\Gamma(r+1)} \int_0^\infty y^{(r+1)-1} e^{-y}\, dy \Big) = \frac{\Gamma(r+1)}{\Gamma(r)} = r$$

(see Exercise 5 for the last step). Hence

$$E(X) = r/\lambda. \tag{20}$$

(b) Suppose again that $X \sim \text{Gamma}(r, \lambda)$ and $Y = \lambda X$. Then

$$E(1/Y) = \int_0^\infty \frac{1}{y} f_Y(y)\, dy = \frac{1}{\Gamma(r)} \int_0^\infty y^{(r-1)-1} e^{-y}\, dy$$
$$= \begin{cases} \Gamma(r-1)/\Gamma(r) = 1/(r-1), & \text{if } r > 1, \\ \infty, & \text{if } r \le 1, \end{cases}$$

and

$$E(1/X) = E(\lambda/Y) = \begin{cases} \lambda/(r-1), & \text{if } r > 1, \\ \infty, & \text{otherwise.} \end{cases} \tag{21}$$

(c) Similar calculations (do them!) show that for $X \sim \text{Beta}(\alpha, \beta)$, with density $x^{\alpha-1}(1-x)^{\beta-1}/B(\alpha, \beta)$ for $0 < x < 1$, one has

$$E(X) = \frac{\alpha}{\alpha + \beta}\,. \tag{22}$$

(d) Suppose $X \sim \chi^2_n = \text{Gamma}(r, \lambda)$ for $r = n/2$ and $\lambda = 1/2$. Then

$$E(X) = \frac{r}{\lambda} = \frac{n/2}{1/2} = n \tag{23}$$

$$E(1/X) = \begin{cases} \lambda/(r-1) = 1/(n-2), & \text{if } n > 2, \\ \infty, & \text{otherwise.} \end{cases} \tag{24}$$

(19) $E(Y_1 Y_2) = E(Y_1)E(Y_2)$ if $Y_1 \ge 0$ and $Y_2 \ge 0$ are independent.

(e) Suppose $X \sim UF(m, n)$. Thus $X = SS_1/SS_2$ where $SS_1 \sim \chi^2_m$ and $SS_2 \sim \chi^2_n$, and SS_1 is independent of SS_2. Since each SS_i is nonnegative, (19) gives

$$E(X) = E(SS_1)E(1/SS_2) = \begin{cases} m/(n-2), & \text{if } n > 2, \\ \infty, & \text{otherwise.} \end{cases} \tag{25}$$

(f) Suppose $X \sim F(m, n)$. Then $X = (SS_1/m)/(SS_2/n) = (n/m)Y$ where $Y \sim UF(m, n)$. Hence

$$E(X) = \frac{n}{m} E(Y) = \begin{cases} n/(n-2), & \text{if } n > 2, \\ \infty, & \text{otherwise.} \end{cases} \tag{26}$$

Example 5. Consider the following game. I am going to pick a number x at random from the F distribution with $m = 3$ and $n = 4$ degrees of freedom. Before I make my draw, you have guess what my x will be; call your guess c. Then I'll make the draw, and you'll pay me

$$(x - c)^2 - w$$

cents (or dollars!), where w is my wager, say 10 units. For example, if you guess my x exactly, I'll pay you 10 units. But if your guess if off by 2, I'll only pay you $10 - 4 = 6$ units, whereas if your guess if off by 4, you'll pay me $16 - 10 = 6$ units. Any takers?

Classroom demonstration here

Some questions: (a) What is the best choice for your guess c? (b) Is it fair for me to wager $w = 10$ units? These questions will be answered in the next lecture. •

We close this section with a couple of simple but useful inequalities.

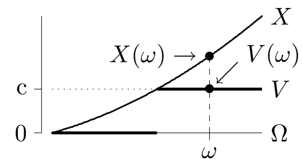
Theorem 5 (Markov's inequality). *Let X be a nonnegative random variable. One has*

$$P[\,X \geq c\,] \leq \frac{E(X)}{c} \tag{27}$$

for each number $c > 0$. Moreover for any given c, equality holds in (27) *if and only if $P[\,X = 0$ or $X = c\,] = 1$.*

Proof Let Ω be the sample space on which X is defined. Let V be the random variable on Ω defined by

$$V(\omega) = \begin{cases} c, & \text{if } X(\omega) \geq c, \\ 0, & \text{otherwise.} \end{cases}$$



Since

$$V(\omega) \leq X(\omega) \tag{28}$$

for all ω, (18) implies that

$$E(V) \leq E(X); \tag{29}$$

(27) follows since

$$E(V) = 0 \times P[\,V = 0\,] + c \times P[\,V = c\,] = c \times P[\,X \geq c\,].$$

If equality holds in (27), then it also holds in (29). By the addendum to (18), equality must hold in (28) for almost all sample points ω, and hence X can take only the values 0 and c, with probability one. Conversely, if X takes just those values, equality does hold in (27). ∎

Theorem 6 (Chebychev's inequality). *Let X be an integrable random variable with mean μ. One has*

$$P[\,|X - \mu| \geq c\,] \leq \frac{E\big((X-\mu)^2\big)}{c^2} \tag{30}$$

for each number $c > 0$. Moreover for any given c, equality holds in (30) *if and only if X takes the values $\mu - c$, μ, and $\mu + c$ with probabilities $(1-p)/2$, p, and $(1-p)/2$ respectively, for some $p \in [0,1]$.*

Chebychev's inequality follows easily from Markov's inequality; the proof is left to you as Exercise 7.

Exercise 1. Let Z be a standard normal random variable. Show that for positive integers k

$$E(Z^k) = \begin{cases} \prod_{j=1}^{k}(2j-1), & \text{if } k = 2j \text{ is even,} \\ 0, & \text{if } k \text{ is odd.} \end{cases} \tag{31}$$ ◇

Exercise 2. Let Y and Z be independent standard normal random variables. For positive integers n, put

$$X_n := Y(1 + Z/\sqrt{n}\,). \tag{32}$$

For $k = 1, 2, \ldots$ find a simple computable expression for $E(X_n^k)$ and show that $E(X_n^k) \to E(Y^k)$ as $n \to \infty$. Evaluate $E(X_n^k)$ for $k = 1$, $\ldots$, 4. ◇

Exercise 3. Let X be an integrable real random variable with distribution function F, quantile function Q, and mean $\mu = E(X)$. Let

$$\delta := E\big(|X - \mu|\big) = \int_{-\infty}^{\infty} |x - \mu|\, F(dx) \tag{33}$$

be the so-called ***mean (absolute) deviation*** (***MAD***) of X about its mean. Show that

$$\begin{aligned} \delta &= \int_0^1 |Q(u) - \mu|\, du \\ &= 2\int_{-\infty}^{\mu} F(x)\, dx = 2\int_{\mu}^{\infty} \big(1 - F(x)\big)\, dx. \end{aligned} \tag{34}$$ ◇

Exercise 4. Show that a random variable X is quasi-integrable if and only if $\sum_{n=1}^{\infty} P[\,|X| \geq n\,] < \infty$. ◇

Exercise 5. Show that the Gamma function $\Gamma(r) := \int_0^\infty x^{r-1}\, e^{-x}\, dx$ satisfies the recursion formula

$$\Gamma(r+1) = r\Gamma(r) \tag{35}$$

for $r > 0$. [Hint: integrate by parts.] ◇

Exercise 6. Find $E(X)$ for random variables X having the following discrete distributions.

Distribution	$P[\,X = k\,]$
Binomial(n, p)	$\binom{n}{k} p^k (1-p)^{n-k},\ k = 0, \ldots, n$
Poisson(μ)	$e^{-\mu}\mu^k/k!\,,\ k = 0, 1, \ldots$
Geometric(p)	$q^{k-1}p,\ k = 0, 1, \ldots$

Exercise 7. Prove Theorem 6. ◇

Exercise 8 (*A weak law of large numbers*). Let $X_1, X_2, \ldots$ be independent random variables, each distributed like a random variable X with $E(X) = 0$ and $\sigma^2 := E(X^2) < \infty$. For each $n \in \mathbb{N}$ set $S_n = X_1 + \cdots + X_n$. (a) Show that $E(S_n) = 0$ and $E(S_n^2) = n\sigma^2$. (b) Show that for each $\epsilon > 0$,

$$\lim_{n\to\infty} P[\,|S_n/n| \geq \epsilon\,] = 0. \tag{36}$$ ◇

Exercise 9. Let $X_1, \ldots, X_n$ be independent random variables, each distributed like a random variable X with $E(X) = 0$ and $E(X^4) < \infty$. (a) Show that X^2 and X^3 are integrable. (b) Put $S_n = X_1 + \cdots + X_n$. Show that

$$E(S_n^4) = nE(X^4) + 3n(n-1)\big(E(X^2)\big)^2. \tag{37}$$

[Hint: Write S_n^4 as $(\sum_{i=1}^n X_i)(\sum_{j=1}^n X_j)(\sum_{k=1}^n X_k)(\sum_{\ell=1}^n X_\ell)$ and expand the sums.] ◇

The following information is needed for next two exercises. Let P be a probability measure on a sample space Ω. Let $A_1, A_2, A_3, \ldots$ be an infinite sequence of events (i.e., subsets of Ω), and let $\limsup_n A_n$ be the set of sample points $\omega \in \Omega$ which belong to A_n for infinitely many n's. According to the ***first Borel-Cantelli lemma***,

$$P[\,\limsup_n A_n\,] = 0 \text{ provided } \textstyle\sum_{n=1}^\infty P[A_n] < \infty. \tag{38}$$

According to the ***second Borel-Cantelli lemma***,

$$P[\,\limsup_n A_n\,] = 1 \text{ provided } \begin{bmatrix} \sum_{n=1}^\infty P[A_n] = \infty \text{ and} \\ \text{the } A_n\text{'s are independent} \end{bmatrix}. \tag{39}$$

Exercise 10. Let $X_1, X_2, \ldots$ be an infinite sequence of independent standard Cauchy random variables and let c be a positive number. Use the second Borel-Cantelli lemma to show that for almost every sample point ω, $X_n(\omega) \geq nc$ for infinitely many n's, and also $X_n(\omega) \leq -nc$ for infinitely many n's. Use this fact to explain the behavior of S_n/n exhibited in Figure 1. ◇

Exercise 11 (*A SLLN*). Let $X_1, X_2, \ldots$ be independent random variables, each distributed like a random variable X with $E(X) = 0$ and $E(X^4) < \infty$. Put $S_n = X_1 + \cdots + X_n$ for each n. Use Markov's inequality for S_n^4, Exercise 9, and the first Borel-Cantelli lemma to show that

$$P[\,|S_n|/n \geq 1/n^{1/8} \text{ for infinitely many } n\,] = 0 \tag{40}$$

and conclude that the set of sample points ω such that $S_n(\omega)/n \to E(X)$ as $n \to \infty$ has probability 1. ◇