EXPECTATION

1. Expectation

Definition 1.1. The expectation, or expected value, or mean value of $X$ with mass function $f$ is defined by

$$E[X] = \sum_x xf(x).$$

whenever the sum is absolutely convergent.

Proposition 1.1. If $X$ has mass function $f$ and $g : \mathbb{R} \to \mathbb{R}$, then

$$E(g(X)) = \sum_x g(x)f(x),$$

whenever this sum is absolutely convergent.

Theorem 1.1. We have

1. If $X \geq 0$, then $E[X] \geq 0$,
2. if $a, b \in \mathbb{R}$, then $E[aX + bY] = aE[X] + bE[Y]$,
3. $E[1\{A\}] = P[A]$,

Proposition 1.2. If $X$ and $Y$ are independent, then

$$E[XY] = E[X]E[Y].$$

2. Variance

Definition 2.1. If $k$ is a positive integer, then the $k$-th moment of $X$ is $E[X^k]$. Moreover the variance $\text{var}(X)$ of $X$ is


Proposition 2.1. For any $a \in \mathbb{R}$,

$$\text{var}(aX) = a^2 \text{var}(X).$$

If $X$ and $Y$ are independent, then

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y).$$
3. USEFUL INEQUALITIES INVOLVING EXPECTATIONS

**Proposition 3.1.** Let $X$ be a random variable taking only non-negative values. Then show that for any $a > 0$, we have
\[ P[X \geq a] \leq \frac{1}{a} E[X]. \]

Here is the famous Cauchy-Schwarz inequality

**Theorem 3.1.** For any $X$ and $Y$
\[ E[XY]^2 \leq E(X^2) E(Y^2). \]

4. USUAL RANDOM VARIABLES

- Bernoulli parameter $p$: $E[B(p)] = p$ and $\text{var}(B(p)) = p(1 - p)$.
- Geometric random variable of parameter $p$: $E[G(p)] = 1/p$ and $\text{var}(G(p)) = (1 - p)/p^2$.
- Binomial of parameter $n$ and $p$: $E[B(n, p)] = np$ and $\text{var}(B(n, p)) = np(1 - p)$.
- Poisson random variable or parameter $\lambda$: $E[P(\lambda)] = \lambda$ and $\text{var}(P(\lambda)) = \lambda$.

5. A FIRST LAW OF LARGE NUMBERS

5.1. A part of the Borel-Cantelli lemma. For a sequence of events $A_n$, we introduce
\[ \{A_n \text{ i.o.}\} = \{A_n \text{ happens infinitely often}\} = \cap_n \cup_{m \geq n} A_m. \]

**Proposition 5.1.** If $\sum_n P(A_n) < \infty$, then
\[ P[A_n \text{ i.o.}] = 0. \]

5.2. Almost sure convergence and strong Law of Large Numbers.

**Definition 5.1.** We say that a $S_n \rightarrow S$ almost surely, if $\{\omega \in \Omega, S_n(\omega) \rightarrow S(\omega)\}$ is an event of probability 1.

We deduce the law of large numbers for i.i.d. random variables with finite fourth moment, i.e. $E[X^4] < \infty$.

**Proposition 5.2.** Let $(X_i)_i$ be a sequence of i.i.d. random variables with $E[X_i^4] < \infty$, we have
\[ \frac{X_1 + \cdots + X_n}{n} \rightarrow E[X], \quad \text{a.s.} \]

**Remark 5.1.** It can be seen by Cauchy-Schwartz that $E[X]^4 \leq E[X^4]$ (take $Y = 1$ in the theorem). It is true in general that $E[X] \leq E[X^p]^{1/p}$ for $p \geq 1$. 
Proofs

Proposition 5.3. If $\sum_n P(A_n) < \infty$, then
$$P[A_n \text{ i.o.}] = 0.$$  

Proof. We have $\{A_k \text{ i.o.}\} \subseteq \bigcup_{m \geq n} A_m$ for any $n$, hence for any $n$
$$P[A_k \text{ i.o.}] \leq P[\bigcup_{m \geq n} A_m] \leq \sum_{m \geq n} P[A_m].$$

By letting $n$ go to infinity the RHS goes to 0, so $P[A_k \text{ i.o.}] = 0$.  \hfill \Box

Proposition 5.4. Let $(X_i)_i$ be a sequence of i.i.d. random variables with $E[X_i^4] < \infty$, we have
$$\frac{X_1 + \cdots + X_n}{n} \to E[X], \quad \text{a.s.}$$

Proof. By substracting $E[X]$ to all $X_i$, we see that $Y_i := X_i - E[X]$ verifies $E[Y_i] = 0$ and $E[Y_i^4] < \infty$. Moreover our proposition is equivalent to
$$\frac{Y_1 + \cdots + Y_n}{n} \to 0, \quad \text{a.s.}$$

Now fix $\varepsilon > 0$ and introduce
$$A_n^\varepsilon = \left\{ \left| \frac{Y_1 + \cdots + Y_n}{n} \right| > \varepsilon \right\}.$$

We have, by Markov’s inequality (Proposition 3.1)
(5.1)  $$P[A_n^\varepsilon] < \frac{1}{(\varepsilon n)^4} E[(Y_1 + \cdots + Y_n)^4].$$

Now
$$E[(Y_1 + \cdots + Y_n)^4] = \sum_i E[Y_i^4] + \sum_{i \neq j} E[Y_i Y_j^3] + \sum_{i \neq j} E[Y_i^2 Y_j^2]$$
$$+ \sum_{i,j,k \text{ distinct}} E[Y_i Y_j Y_k^2] + \sum_{i,j,k,l \text{ distinct}} E[Y_i Y_j Y_k Y_l],$$

which considering the fact that the $Y_i$’s are i.i.d. and with 0 mean we see that
$$E[(Y_1 + \cdots + Y_n)^4] = \sum_i E[Y_i^4] + \sum_{i \neq j} E[Y_i^2] E[Y_j^2],$$

(e.g. $\sum_{i \neq j} E[Y_i Y_j^3] = \sum_{i \neq j} E[Y_i^2] E[Y_j^3] = 0$ by Proposition 1.2).

Now recalling that $E[Y_i^4] < \infty$ implies that $E[Y_i^2] < \infty$, we can see that there exists a constant $C < \infty$ such that
$$E[(Y_1 + \cdots + Y_n)^4] \leq Cn^2,$$
which using (5.1) means that

\[ P[A_n^\varepsilon] < C \frac{1}{\varepsilon^4 n^2}, \]

which is summable. Using proposition 5.3, this implies that for any \( \varepsilon > 0 \)

\[ P[A_n^\varepsilon \text{ i.o.}] = 0. \]

Considering the special case where \( \varepsilon = 1/k \) for \( k \in \mathbb{N} \) we see

\[ P[A_n^{1/k} \text{ i.o.}] = 0, \]

such that taking a countable union

\[ P[\bigcup_{k \geq 0} A_n^{1/k} \text{ i.o.}] = 0. \]

and by taking the complement (and recalling the definition of \( \{A_n^{1/k} \text{ i.o.}\} \)), we see that

\[ P\left[ \bigcap_{k \geq 0} \bigcup_{n \geq 0} \bigcap_{m \geq n} \left\{ \frac{Y_1 + \cdots + Y_n}{n} < 1/k \right\} \right] = 1. \]

Hence the set \( \cap_{k \geq 0} \cup_{n \geq 0} \cap_{m \geq n} \{ |Y_1 + \cdots + Y_n| < 1/k \} \) has probability 1. Now let us see what this horrible thing means, if \( \omega \in \cap_{k \geq 0} \cup_{n \geq 0} \cap_{m \geq n} \{ |Y_1 + \cdots + Y_n| < 1/k \} \) we have

\[ \forall k > 0, \exists n \geq 0, \forall m \geq n, \quad \left| \frac{Y_1(\omega) + \cdots + Y_n(\omega)}{n} \right| < 1/k, \]

which exactly means that for \( \omega \in \cap_{k \geq 0} \cup_{n \geq 0} \cap_{m \geq n} \{ |Y_1 + \cdots + Y_n| < 1/k \} \)

\[ \frac{Y_1(\omega) + \cdots + Y_n(\omega)}{n} \to 0. \]

As we argued at the beginning of the proof, this is enough to prove our proposition. \( \square \)