1. If $X_n \geq 0$, $X_n \downarrow X$ a.s. and $EX_N < \infty$ for some $N$, then $EX_n \to EX$

**proof** Consider the variable $Y_n = X_N - X_n$, for $n \geq N$. Since $X_n$ is decreasing, $(Y_n)_{n \geq N}$ is non-negative increasing, and its limit is $X_N - X$. Therefore by the MCT, we have

$$\lim_{n \to \infty} E(Y_n) = EX_N - EX$$

and using linearity of the expectation,

$$\lim_{n \to \infty} E(X_N) - E(X_n) = EX_N - EX$$

finally, since $EX_N < \infty$, we have

$$\lim_{n \to \infty} E(X_n) = EX$$

2. if $E|X| < \infty$ then $E|X|1_{(|X|>n)} \to 0$ as $n \to \infty$

**proof** Let $A = \{\omega : |X(\omega)| < \infty\}$. Since $E|X| < \infty$, we have $\mu(A^c) = 0$, and we can restrict the integral to $A^c$:

$$E|X| = \int_{\Omega} |X(\omega)|1_A(\omega)d\mu(\omega)$$

now consider the sequence of functions

$$f_n(\omega) = |X(\omega)|1_{|X|>n}(\omega)$$

Then

- $f_n$ is non-negative, measurable
- for all $\omega$, $f_n(\omega) \uparrow |X(\omega)|1_A$ as $n \to \infty$

Thus by the MCT, we have $\lim_{n \to \infty} E f_n(\omega) = EX$, which is the desired result.

3. if $E|X_1| < \infty$ and $X_n \uparrow X$ a.s., then either $EX_n \uparrow EX < \infty$ or else $EX_n \uparrow \infty$ and $E|X| = \infty$
First, since $X_1$ is integrable, we have $E X_1^+ < \infty$ and $E X_1^- < \infty$. Now since $X_n \uparrow X$, we have $X_n^+ \uparrow X^+$ and $X_n^- \downarrow X^-$, where $X_n^+$ and $X_n^-$ are non-negative. We have

- Since $E X_1^- < \infty$, all the terms of the sequence are finite, and we have (by the MCT) $E X_n^- \downarrow E X^-$, which is finite.
- By the MCT, $E X_n^+ \uparrow E X^+$. Two cases are possible
  
(a) either the limit $E X^+ = \infty$, in which case $E X_n = E X_n^+ - E X^-$ is the sum of a sequence that diverges to $\infty$ and a bounded sequence, therefore it diverges, and $E X = E X^+ + E X^- = \infty$.

(b) either the limit $E X^+ < \infty$, in which case $E X_n = E X_n^+ - E X^-$ is the sum of two converging sequences, thus it converges to the sum of limits, i.e. $E X_n \rightarrow E X < \infty$.

4. if $X$ takes values in the non-negative integers, then

$$E X = \sum_{n=1}^{\infty} P(X \geq n)$$

By definition, the expectation of $X$ is given by

$$E X = \sum_{i=1}^{\infty} i P(X = i)$$

which is the sum of the series with terms $f_i = i P(X = i)$. Now consider the sequence of series $f^{(m)}$, $m \geq 1$, where each $f^{(m)}$ is given by

$$f^{(m)}_i = \sum_{n=1}^{m} P(X = i) 1_{n \leq i}$$

then we have

- $f^{(m)}$ is non-negative, measurable for all $m$
- for all $i$, $\lim_{m \rightarrow \infty} f^{(m)}_i = \sum_{n \leq i} P(X = i) = i P(X = i) = f_i$

Therefore by the MCT, we have

$$\sum_{i=1}^{\infty} f_i = \lim_{m \rightarrow \infty} \sum_{i=1}^{m} f^{(m)}_i$$

now

$$\sum_{i=1}^{\infty} f^{(m)}_i = \sum_{i=1}^{m} \sum_{n=1}^{\infty} P(X = i) 1_{n \leq i}$$

$$= \sum_{n=1}^{m} \sum_{i=1}^{\infty} P(X = i) 1_{n \leq i} \quad \text{by linearity of the integral}$$

$$= \sum_{n=1}^{m} P(X \geq n)$$

therefore we have

$$E X = \lim_{m \rightarrow \infty} \sum_{n=1}^{m} P(X \geq n)$$

which is the desired result
1. For a counting r.v. $X = \sum_{i=1}^{n} 1_{A_i}$, give a formula for the variance of $X$ in terms of the probabilities $P(A_i)$ and $P(A_i \cap P_j), i \neq j$.

**answer** We have by linearity of the expectation,

$$E(X) = \sum_{i=1}^{n} P(A_i)$$

then

$$E(X^2) = E(\sum_{i=1}^{n} \sum_{j=1}^{n} 1_{A_i} 1_{A_j})$$

$$= \sum_{i=1}^{n} E(1_{A_i}) + 2 \sum_{i<j} E(1_{A_i \cap A_j})$$

$$= \sum_{i=1}^{n} P(A_i) + 2 \sum_{i<j} P(A_i \cap A_j)$$

therefore the variance is

$$\text{var}(X) = E(X^2) - (E(X))^2$$

$$= \sum_{i=1}^{n} P(A_i) + 2 \sum_{i<j} P(A_i \cap A_j) - \left(\sum_{i=1}^{n} P(A_i)\right)^2$$

2. If $k$ balls are put at random into $n$ boxes, what is the variance of $X = \text{number of empty boxes}$?

**answer** We can write $X$ as the sum of indicator variables

$$X = \sum_{i=1}^{k} 1_{A_i}$$

where $A_i$ is the event “box $i$ is empty”. Here we have for all $i$, $P(A_i) = (1 - 1/k)^n$ (all $n$ balls should be put in any box other than the $i$-th one) and similarly for all $i \neq j$, $P(A_i \cap A_j) = (1 - 2/k)^n$. Applying the above expression, we obtain an expression for the variance

$$\text{var}(X) = \sum_{i=1}^{n} P(A_i) + 2 \sum_{i<j} P(A_i \cap A_j) - \left(\sum_{i=1}^{n} P(A_i)\right)^2$$

$$= n(1 - 1/k)^n + 2n(n - 1)(1 - 2/k)^n - n^2(1 - 1/k)^{2n}$$

(3.3)

1. Suppose $E(X) = 0$ and $\text{var}(X) = \sigma^2 < \infty$. Prove

$$P(X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$
**Proof** Let $c > 0$, and consider the function $\phi : x \mapsto (x + c)^2$. $\phi$ is non-negative, increasing on $(a, \infty)$ (since $a > 0 > -c$).

We have by Chebyshev’s inequality

$$P(X \geq a) \leq \frac{E \phi(X)}{\phi(a)} = \frac{E X^2 + 2cEX + c^2}{(a + c)^2} = \frac{\sigma^2 + c^2}{(a + c)^2}$$

minimizing the bound over $c$, we have

$$P(X \geq a) \leq \frac{\sigma^2 + c^2}{(a + c)^2}$$

2. Suppose $X \geq 0$ and $EX^2 < \infty$. Prove

$$P(X > 0) \geq \frac{(EX)^2}{EX^2}$$

**Proof** Applying Cauchy-Schwarz inequality to the variables $X$ and $1_{(X>0)}$, we have

$$E(X1_{(X>0)}) \leq \sqrt{E(X^2)E1^2_{(X>0)}}$$

or, squaring both sides,

$$(EX)^2 \leq E(X^2)P(X > 0)$$

which is the desired inequality.

(3.4 Chebyshev’s other inequality) Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be bounded and increasing functions. Prove that for any r.v. $X$,

$$E(f(X)g(X)) \geq (E f(X))(E(g(X)))$$

in other words, $f(X)$ and $g(X)$ are positively correlated.

**Proof** Using the definition of the expectation, we can write

$$E f(X)E g(X) = \int_{\Omega} f(X(\omega_1))d\mu(\omega_1) \int_{\Omega} g(X(\omega_2))d\mu(\omega_2) = \iint f(X(\omega_1))g(X(\omega_2))d\mu(\omega_1)d\mu(\omega_2)$$

By Fubini’s theorem

now partitioning $\Omega \times \Omega = A \cup A^c$, where $A = \{(\omega_1, \omega_2) \in \Omega \times \Omega : X(\omega_1) \leq X(\omega_2)\}$, and using the fact that on $A$, $f(X(\omega_1)) \leq f(X(\omega_2))$ and on $A^c$, $g(X(\omega_2)) \leq g(X(\omega_1))$ ($f$ and $g$ are increasing)

$$E f(X)E g(X) = \iint f(X(\omega_1))g(X(\omega_2))1_A(\omega_1, \omega_2)d\mu(\omega_1)d\mu(\omega_2) + \int f(X(\omega_1))g(X(\omega_2))1_{A^c}(\omega_1, \omega_2)d\mu(\omega_1)d\mu(\omega_2) \leq \iint f(X(\omega_2))g(X(\omega_2))1_A(\omega_1, \omega_2)d\mu(\omega_1)d\mu(\omega_2) + \int f(X(\omega_1))g(X(\omega_1))1_{A^c}(\omega_1, \omega_2)d\mu(\omega_1)d\mu(\omega_2)$$

$$= \iint f(X(\omega_2))g(X(\omega_2))h_2(\omega_2)d\mu(\omega_2) + \int f(X(\omega_1))g(X(\omega_1))h_1(\omega_1)d\mu(\omega_1)$$
where \( h_1(\omega_1) = \int_\Omega 1_A(\omega_1, \omega_2) d\mu(\omega_2) = P(X(\omega_1) \leq X), \) and \( h_2(\omega_2) = P(X \leq X(\omega_2)) \), in particular, for all \( \omega, \ h_1(\omega) + h_2(\omega) = 1. \) Therefore we have

\[
E f(X) E g(X) \leq \int f(X(\omega_2)) g(X(\omega_2)) h_2(\omega_2) d\mu(\omega_2) + \int f(X(\omega_1)) g(X(\omega_1)) h_1(\omega_1) d\mu(\omega_1)
\]

by renaming the integration variables

\[
= \int f(X(\omega)) g(X(\omega)) (h_2(\omega) + h_1(\omega)) d\mu(\omega)
\]

which concludes the proof.

(3.5) Let \( X \sim \text{Poisson}(\lambda) \) and \( Y \sim \text{Poisson}(2\lambda). \)

1. Prove \( P(X \geq Y) \leq \exp(-(3 - \sqrt{8})\lambda) \) if \( X \) and \( Y \) are independent.

**proof** Using a Chernoff type bound, we have for all \( \alpha > 0 \)

\[
P(X \geq Y) = P(\alpha^{X-Y} \geq 1)
\]

\[
\leq E(\alpha^{X-Y}) \quad \text{by Markov's inequality}
\]

\[
= E\alpha^X E\frac{1}{\alpha}^Y \quad \text{by independence}
\]

where

\[
E\alpha^X = \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\alpha^k = \exp(-\lambda) \exp(\lambda\alpha)
\]

and similarly

\[
E\frac{1}{\alpha}^Y = \exp(-2\lambda) \exp\left(\frac{2\lambda}{\alpha}\right)
\]

therefore

\[
P(X \geq Y) \leq \exp(\lambda(-3 + \alpha + \frac{2}{\alpha}))
\]

for all \( \alpha > 0. \) Let \( b(\alpha) = \exp(\lambda(-3 + \alpha + \frac{2}{\alpha})) \) be the upper bound. Optimizing the bound over \( \alpha, \) we have

\[
b'(\alpha) = 0 \iff 1 - \frac{2}{\alpha^2} = 0
\]

\[
\iff \alpha = \sqrt{2}
\]

thus for \( \alpha = \sqrt{2}, \) we obtain the desired bound

\[
P(X \geq Y) \leq \exp(\lambda(-3 + 2\sqrt{2}))
\]

2. Find constants \( A < \infty, \ c > 0, \) not depending on \( \lambda, \) such that without assuming independence, \( (X \geq Y) \leq A \exp(-c\lambda) \)
Since we do not assume independence, we can use Cauchy-Schwarz inequality to bound the expectation of the product. For all $\alpha > 0$, we have

$$P(X \geq Y) = P(\alpha X - Y \geq 1)$$

$$\leq E(\alpha X - Y) \quad \text{by Markov’s inequality}$$

$$\leq \sqrt{E(\alpha X)^2 E \left(\frac{Y}{\alpha}\right)^2} \quad \text{by Cauchy-Schwarz}$$

$$\leq \sqrt{E(\alpha^2 X) E \frac{Y}{\alpha^2}}$$

The product of expectations is bounded by $b(\alpha^2)$, which is minimal for $\alpha^2 = \sqrt{2}$, and we obtain the following bound

$$P(X \geq Y) \leq \sqrt{\exp(\lambda(-3 + 2\sqrt{2}))} = \exp(\lambda(-\frac{3}{2} + \sqrt{2}))$$

we obtain the desired bound, where $A = 1$ and $c = \frac{3}{2} - \sqrt{2} > 0$. 
