Chebyshev’s Inequality

Let \( X \) be a random variable taking nonnegative values and having a finite expectation. For simplicity, we will assume that \( X \) is discrete (the general case is treated similarly).

Let \( x_i \geq 0 \) be the values of \( X \), and consider the mean value
\[
E[X] = \sum_i x_i P(X = x_i).
\]
Every term in the above sum is nonnegative. If we omit the terms with \( x_i \) below a fixed bound \( t > 0 \), the net value of the sum would not be increased:
\[
\sum_i x_i P(X = x_i) \geq \sum_{i: x_i \geq t} x_i P(X = x_i).
\]
Observe that
\[
\sum_{i: x_i \geq t} x_i P(X = x_i) \geq \sum_{i: x_i \geq t} tP(X = x_i) = tP(X \geq t).
\]
Therefore, putting the chain of inequalities together,
\[
P(X \geq t) \leq \frac{E[X]}{t}, \quad t > 0.
\]
Actually, this estimate is only meaningful for \( t > E[X] \). We may also write it in the form
\[
P(X < t) \geq 1 - \frac{E[X]}{t}.
\]
Let now \( Y \) be a (real) random variable without value restrictions, and let \( \mu \) be its mean. Then \( X = (Y - \mu)^2 \) is nonnegative. So, according to the derived estimate,
\[
P((Y - \mu)^2 \geq t) \leq \frac{E[(Y - \mu)^2]}{t}, \quad t > 0.
\]
Letting \( s = \sqrt{t} \), and noting that \((Y - \mu)^2 \geq t \) is equivalent to \( |Y - \mu| \geq s \), we obtain
\[
P(|Y - \mu| \geq s) \leq \frac{\text{Var}(Y)}{s^2}.
\]
This is Chebyshev’s inequality. It is meaningful for any \( Y \) with finite mean and variance and \( s > \sigma \), where \( \sigma \) is the standard deviation of \( Y \).

For the complementary event \( |Y - \mu| < s \), we have \( P(|Y - \mu| < s) \geq 1 - \frac{\text{Var}(Y)}{s^2} \).

A restatement of Chebyshev’s inequality is often in use. If \( s = k\sigma \), the inequality becomes
\[
P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}.
\]
For instance, the probability that \( Y \) falls two standard deviations away from its mean is under \( 1/4 \).