Large deviation bounds summary

1. Markov's inequality: $\mathrm{P}(X \geq a) \leq \mathrm{E}[X]/a$.
   - Applies to any *non-negative* random variable X, and any $a > 0$ ($a > \mathrm{E}[X]$ for a meaningful bound).
   - Requires only knowledge of $\mathrm{E}[X]$
   - Generally useful when $\mathrm{E}[X]$ is small and X is "concentrated" around $\mathrm{E}[X]$.
2. Chebyshev's inequality: $\mathrm{P}(|X - \mu| \geq t\sigma) \leq 1/t^2$, where $\mu = \mathrm{E}[X], \sigma = \sqrt{\mathrm{Var}[X]}$.
   - Applies to any random variable X (with finite μ, σ), and any $t > 0$ ($t > 1$ for a meaningful bound).
   - Requires knowledge of both $\mathrm{E}[X]$ and $\mathrm{Var}[X]$.
   - Can be used to bound both $\mathrm{P}(X \geq a)$ and $\mathrm{P}(X \leq a)$.
3. Central Limit Theorem: If $X = X_1 + \cdots + X_n$ where X_i independent and have same PDF/PMF, then $(X - \mathrm{E}[X])/\sqrt{\mathrm{Var}[X]} \approx \mathrm{Normal}(0, 1)$.
   - Applies for X being sum of many *independent* random variables.
   - Requires knowledge of $\mu = \mathrm{E}[X_i]$ and $\sigma^2 = \mathrm{Var}[X_i]$ to obtain $\mathrm{E}[X] = n\mu$, $\mathrm{Var}[X] = n\sigma^2$.
   - Approximates the CDF of X, but does not provide an error on the quality of the approximation.[1] Using the axioms, we can use it to approximate probabilities of other events like $\mathrm{P}(150 < |X| < 200)$.

[1] This error will depend on the PDF/PMF of X_i. The Berry-Esseen Theorem is a refinement of the Central Limit Theorem that gives a explicit error bound. Away from the mean, Chernoff bounds give much tighter estimates for many random variables of interest.