<u>Markov's Inequality</u>: Let X be a random variable that assumes only nonnegative values. Then, for all $a > 0$, $Pr(X \geq a) \leq \frac{E(X)}{a}$.

For $a > 0$, let $I = \begin{cases} 1 & \text{if } X \geq a, \\ 0 & \text{otherwise}, \end{cases}$

and note that, since $X \geq 0$, $I \leq \frac{X}{a} \longrightarrow (1)$

Because I is a 0-1 random variable,
$E(I) = Pr(I=1) = Pr(X \geq a)$.

Taking expectations in (1) gives:
$$Pr(X \geq a) = E(I) \leq E\left(\frac{X}{a}\right) = \frac{E(X)}{a}.$$

The <u>Variance</u> of a random variable X is defined as:

$$Var[X] = E[(X - E(X))^2] = E[X^2 - 2XE(X) + E(X)^2]$$
$$= E(X^2) - 2E[XE(X)] + E(X)^2$$
$$= E(X^2) - 2E(X)E(X) + E(X)^2$$
$$= E(X^2) - (E(X))^2$$

Here we have used linearity of expectation.

If X and Y are two independent random variables, then $E[X \cdot Y] = E(X) \cdot E(Y)$.

$$E[X \cdot Y] = \sum_i \sum_j (i \cdot j) \cdot \Pr((X=i) \cap (Y=j))$$

$$= \sum_i \sum_j (i \cdot j) \cdot \Pr(X=i) \cdot \Pr(Y=j)$$

$$= \left(\sum_i i \cdot \Pr(X=i)\right)\left(\sum_j j \cdot \Pr(Y=j)\right)$$

$$= E(X) \cdot E(Y).$$

<u>Chebyshev's Inequality:</u> For any $a > 0$,

$$\Pr(|X - E(X)| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}$$

$$\Pr(|X - E(X)| \geq a) = \Pr((X - E(X))^2 \geq a^2).$$

Applying Markov's inequality:

$$\Pr((X-E(X))^2 \geq a^2) \leq \frac{E[(X-E(X))^2]}{a^2} = \frac{\mathrm{Var}[X]}{a^2}$$

<u>Chernoff Bounds for the Sum of Poisson Trials:</u>

The distributions of the random variables in Poisson trials are not necessarily identical. Bernoulli trials are a special case of Poisson trials where the independent 0-1 random variables have the same distribution.

Let X_i $(1 \leq i \leq n)$ be independent random variables with $\Pr(X_i = 1) = p_i$. Let $X = \sum_{i=1}^{n} X_i$, and let

$$\mu = E(X) = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p_i$$

For a given $\delta > 0$, we are interested in bounds on $\Pr(X \geq (1+\delta)\mu)$ and $\Pr(X \leq (1-\delta)\mu)$ – that is, the probability that X deviates from its expectation μ by $\delta\mu$ or more.

$$E[e^{tX_i}] = p_i e^t + (1-p_i) = 1 + p_i(e^t - 1) \leq e^{p_i(e^t-1)}$$

$$\Rightarrow E[e^{tX}] = \prod_{i=1}^{n} E[e^{tX_i}] \leq e^{(e^t-1)\sum_{i=1}^{n} p_i} = e^{(e^t-1)\mu}$$

For any $t > 0$: $\Pr(X \geq a) = \Pr(e^{tX} \geq e^{ta}) \leq \dfrac{E[e^{tX}]}{e^{ta}}$

$\leq \dfrac{e^{(e^t-1)\mu}}{e^{ta}}$. Putting $a = (1+\delta)\mu$, we get:

$$\Pr[X \geq (1+\delta)\mu] \leq \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}$$

For any $\delta > 0$, we can set $t = \log_e(1+\delta) > 0$ to get:

$$\Pr(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

For any $t < 0$: $\Pr[X \leq (1-\delta)\mu] = \Pr[e^{tX} \geq e^{(1-\delta)\mu t}]$

$\leq \dfrac{E[e^{tX}]}{e^{(1-\delta)\mu t}} \leq \dfrac{e^{(e^t-1)\mu}}{e^{(1-\delta)\mu t}}$. For $0 < \delta < 1$, we set $t = \log_e(1-\delta) < 0$ to get:

$$\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$