CMPT 210: Probability and Computing

Lecture 22

Sharan Vaswani

April 2, 2024

Recap

Tail inequalities bound the probability that the r.v. takes a value much different from its mean.

Markov's Theorem: If X is a non-negative random variable, then for all $x > 0$, $\Pr[X \geq x] \leq \frac{\mathbb{E}[X]}{x}$.

Chebyshev's Theorem: For a r.v. X and all $x > 0$, $\Pr[|X - \mathbb{E}[X]| \geq x] \leq \frac{\text{Var}[X]}{x^2}$.

Pairwise Independent Sampling

Claim: Let $G_1, G_2, \dots, G_n$ be pairwise independent random variables with the same mean μ and standard deviation σ. Define $S_n := \sum_{i=1}^{n} G_i$, then,

$$\Pr\left[\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right] \leq \frac{1}{n}\left(\frac{\sigma}{\epsilon}\right)^2.$$

Proof: Let us compute $\mathbb{E}[S_n/n]$ and $\mathrm{Var}[S_n/n]$.

$$\mathbb{E}[S_n] = \mathbb{E}\left[\sum_{i=1}^{n} G_i\right] = \sum_{i=1}^{n} \mathbb{E}[G_i] = n\mu \implies \mathbb{E}[S_n/n] = \frac{1}{n}\mathbb{E}[S_n] = \mu$$

(Using linearity of expectation)

$$\mathrm{Var}[S_n] = \mathrm{Var}\left[\sum_{i=1}^{n} G_i\right] = \sum_{i=1}^{n} \mathrm{Var}[G_i] = n\sigma^2$$

(Using linearity of variance for pairwise independent r.v's)

$$\implies \mathrm{Var}[S_n/n] = \frac{1}{n^2}\mathrm{Var}[S_n] = \frac{\sigma^2}{n}$$

Pairwise Independent Sampling

Using Chebyshev's Theorem,

$$\Pr\left[\left|\frac{S_n}{n} - \mathbb{E}\left[\frac{S_n}{n}\right]\right| \geq \epsilon\right] = \Pr\left[\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right] \leq \frac{\mathsf{Var}[S_n/n]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

Hence, for arbitrary pairwise independent r.v's, if n increases, the probability of deviation from the mean μ decreases.

Weak Law of Large Numbers: Let $G_1, G_2, \ldots, G_n$ be pairwise independent variables with the same mean μ and (finite) standard deviation σ. Define $X_n := \frac{\sum_{i=1}^{n} G_i}{n}$, then for every $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr[|X_n - \mu| \leq \epsilon] = 1.$$

Proof: Follows from the theorem on pairwise independent sampling since $\lim_{n\to\infty} \Pr[|X_n - \mu| \leq \epsilon] = \lim_{n\to\infty}\left[1 - \frac{\sigma^2}{n\epsilon^2}\right] = 1$.

Questions?

Sums of Random Variables

If we know that the r.v X is (i) non-negative and (ii) $\mathbb{E}[X]$, we can use Markov's Theorem to bound the probability of deviation from the mean.

If we know both (i) $\mathbb{E}[X]$ and (ii) $\mathrm{Var}[X]$, we can use Chebyshev's Theorem to bound the probability of deviation.

In many cases the random variable of interest is a sum of r.v's (e.g., for the voter poll application), and we can use the Chernoff bound to obtain tighter bounds on the deviation from the mean.

Chernoff Bound: Let $T_1, T_2, \ldots, T_n$ be mutually independent r.v's such that $0 \leq T_i \leq 1$ for all i. If $T := \sum_{i=1}^{n} T_i$, for all $c \geq 1$ and $\beta(c) := c \ln(c) - c + 1$,

$$\Pr[T \geq c\mathbb{E}[T]] \leq \exp(-\beta(c)\,\mathbb{E}[T])$$

If $T_i \sim \mathrm{Ber}(p)$ and are mutually independent, then $T_i \in \{0, 1\}$ and we can use the Chernoff bound to bound the deviation from the mean for $T \sim \mathrm{Bin}(n, p)$. In general, if $T_i \in [0, 1]$, the Chernoff Bound can be used even if the T_i's have different distributions!

Chernoff Bound – Binomial Distribution

Q: Bound the probability that the number of heads that come up in 1000 independent tosses of a fair coin exceeds the expectation by 20% or more.

Let T_i be the indicator r.v. for the event that coin i comes up heads, and let T denote the total number of heads. Hence, $T = \sum_{i=1}^{1000} T_i$. For all i, $T_i \in \{0, 1\}$ and are mutually independent r.v's. Hence, we can use the Chernoff Bound.

We want to compute the probability that the number of heads is larger than the expectation by 20% meaning that $c = 1.2$ for the Chernoff Bound. Computing $\beta(c) = c\ln(c) - c + 1 \approx 0.0187$. Since the coin is fair, $\mathbb{E}[T] = 1000\,\frac{1}{2} = 500$. Plugging into the Chernoff Bound,

$$\Pr[T \geq c\mathbb{E}[T]] \leq \exp(-\beta(c)\mathbb{E}[T]) \implies \Pr[T \geq 1.2\,\mathbb{E}[T]] \leq \exp(-(0.0187)\,(500)) \approx 0.0000834.$$

Comparing this to using Chebyshev's inequality,

$$\begin{aligned}\Pr[T \geq c\mathbb{E}[T]] = \Pr[T - \mathbb{E}[T] \geq (c-1)\mathbb{E}[T]] &\leq \Pr[|T - \mathbb{E}[T]| \geq (c-1)\,\mathbb{E}[T]] \\ &\leq \frac{\mathrm{Var}[T]}{(c-1)^2\,(\mathbb{E}[T])^2} = \frac{1000\,\frac{1}{4}}{(1.2-1)^2(500^2)} = \frac{250}{0.2^2\,500^2} = \frac{250}{10000} = 0.025.\end{aligned}$$