Lecture 05: Concentration Bounds
Markov Inequality

Theorem (Markov Inequality)

Let $X$ be a positive random variable. Then the following holds:

$$P[X \geq t] \leq \frac{E[X]}{t}$$

- We only present the proof when the sample space $\Omega$ is discrete.
- Suppose this statement is false, i.e.,

$$P[X \geq t] > \frac{E[X]}{t}$$
Then we can perform the following analysis

\[
\mathbb{E}[X] = \sum_{i \in \Omega} i \cdot \mathbb{P}[X = i]
\]

\[
= \sum_{i \in \Omega: i < t} i \cdot \mathbb{P}[X = i] + \sum_{i \in \Omega: i \geq t} i \cdot \mathbb{P}[X = i]
\]

\[
\geq \sum_{i \in \Omega: i < t} 0 \cdot \mathbb{P}[X = i] + \sum_{i \in \Omega: i \geq t} t \cdot \mathbb{P}[X = i]
\]

\[
= t \cdot \mathbb{P}[X \geq t] > \mathbb{E}[X]
\]

Hence, contradiction. This proves the theorem.

The probability distribution that shows that this inequality is tight is:

\[
\mathbb{P}[X = 0] = 1 - 1/t \quad \text{and} \quad \mathbb{P}[X = t] = 1/t
\]
Chebyshev’s Inequality

Theorem (Chebyshev’s inequality)

\[ P \left[ |X - \mathbb{E}[X]| \geq t \right] \leq \frac{\text{Var}[X]}{t^2} \]

- Note that \( P \left[ |X - \mathbb{E}[X]| \geq t \right] = P \left[ (X - \mathbb{E}[X])^2 \geq t^2 \right] \)
- Apply Markov Inequality
Some Exercises

- Compute the $E[X]$ and $Var[X]$ for the following probability distributions
  1. $P[X = 0] = 1 - 1/t$ and $P[X = t] = 1/t$
  2. For positive constant $p$, $P[X = 0] = 1 - p$ and $P[X = 1] = p$.

- Prove the following properties for independent probability distributions $X$ and $Y$
  1. Prove that $E[exp(X + Y)] = E[exp(X)] \cdot E[exp(Y)]$
  2. Prove that $Var[X + Y] = Var[X] + Var[Y]$
Let $X$ be the probability distribution such that $\mathbb{P}[X = 0] = 1 - p$ and $\mathbb{P}[X = 1] = p$ (Bernoulli variable).

Let $B(n, p)$ be the random variable $\sum_{i=1}^{n} X(i)$, where $X(i)$ is the $i$-th independent sample according to the distribution $X$.

$B(n, p)$ (sum of $n$ independent Bernoulli variables)

**Theorem (Chernoff Bound)**

For $0 < t < 1 - p$, we have

$$\mathbb{P} \left[ B(n, p) - np \geq nt \right] \leq 2^{-nD_{KL}(p+t,p)} \leq \exp \left( -2t^2 n \right)$$

$D_{KL}(\cdot, \cdot)$ is the Kullback-Leibler divergence and is defined as follows:

$$D_{KL}(\alpha, \beta) := \alpha \log \left( \frac{\alpha}{\beta} \right) + (1 - \alpha) \log \left( \frac{1 - \alpha}{1 - \beta} \right)$$

It is always $\geq 0$. Why?
Comments on the Bound

- Substituting \( t = 1 / \sqrt{n} \), we get that the probability
  \[ \Pr[ B(n, p) \geq np + \sqrt{n} ] \leq \text{const.}, \]
  \( B(n, p) \) is very strongly concentrated around its mean!

- Chernoff Bound is also very tight. That is, we have
  \[ \Pr[ B(n, p) - np \geq nt ] \geq \frac{\exp(-2t^2n)}{\text{poly}(n)} \]
  This can be proven using “Stirling Approximation” or “The Method of Types”

- Even using limited independence, one can get the same bound as Chernoff Bound: “Chernoff-Hoeffding Bounds for Applications with Limited Independence,” by Jeanette P. Schmidt, Alan Siegel, and Aravind Srinivasan.
Proof of Chernoff Bound

- We are interested in bounding
  \[ P \left[ B(n, p) \geq n(p + t) \right] \]

- For any \( h > 0 \), this probability is identical to
  \[ P \left[ \exp \left( hB(n, p) \right) \geq \exp \left( hn(p + t) \right) \right] \]
By Markov Inequality, we get

\[
\mathbb{E} \left[ \exp \left( hB(n, p) \right) \right] \leq \frac{\exp \left( hn(p + t) \right)}{\exp \left( hn(p + t) \right)} \]

\[
= \mathbb{E} \left[ \exp \left( h \sum_{i=1}^{n} X^{(i)} \right) \right] \frac{\exp \left( hn(p + t) \right)}{\exp \left( hn(p + t) \right)} \]

\[
= \prod_{i=1}^{n} \mathbb{E} \left[ \exp \left( hX^{(i)} \right) \right] \frac{\exp \left( hn(p + t) \right)}{\exp \left( hn(p + t) \right)} \]

\[
= \left( \frac{1 - p + p \exp(h)}{\exp \left( h(p + t) \right)} \right)^n
\]
This expression is an upper-bound for all $h > 0$.

So, we choose $h$ that minimizes the upper bound expression.

Using basic calculus, it can be shown that the expression

$$E = \frac{1 - p + p \exp(h)}{\exp(h(p + t))}$$

is minimized for

$$\exp(h) = \frac{(1 - p)(p + t)}{p(1 - p - t)}$$

Note that here we are using the fact that $(1 - p - t) > 0$.
Now, let us calculate the expression $E$

\[
1 - p + \frac{(1-p)(p+t)}{(1-p)(p+t)} = \frac{(1-p) p^{p+t} (1-p-t)^{p+t}}{(1-p-t) (1-p)^{p+t} (p+t)^{p+t}}
\]

\[
= \left( \frac{1-p}{1-p-t} \right)^{1-p-t} \cdot \left( \frac{p}{p+t} \right)^{p+t}
\]

Note that the expression $E$ we obtained satisfies

\[-\log E = D_{KL} (p + t, p)\]

Therefore we get the first part of the Chernoff bound:

\[
\mathbb{P} \left[ B(n, p) \geq n(p + t) \right] \leq E^n = 2^{-nD_{KL}(p+t,p)}
\]
For the final part of the inequality, we need to show that

\[
\left( \frac{1 - p}{1 - p - t} \right)^{1-p-t} \cdot \left( \frac{p}{p + t} \right)^{p+t} \leq \exp \left( -2t^2 \right)
\]
Proof of Chernoff Bound

- First Attempt (Which will fail)

\[
\frac{1 - p}{1 - p - t} = 1 + \frac{t}{1 - p - t} \leq \exp \left( \frac{t}{1 - p - t} \right)
\]

Using the fact that \(1 + x \leq \exp(x)\)

\[
\frac{p}{p + t} = 1 - \frac{t}{p + t} \leq \exp \left( \frac{t}{p + t} \right)
\]

Using the fact that \(1 - x \leq \exp(-x)\)

Therefore,

\[
E = \left( \frac{1 - p}{1 - p - t} \right)^{1-p-t} \cdot \left( \frac{p}{p + t} \right)^{p+t}
\]

\[
\leq \exp(t) \cdot \exp(-t) = 1
\]

- Although this is true, but we do not get anything nontrivial.
- Our target is to show that \(E \leq \exp(-2t^2)\).
Proof of Chernoff Bound

- Second Attempt
- We use $\bar{p} = 1 - p$ for brevity
- $f(t) := \ln E = (p + t) \ln \left( \frac{p}{p+t} \right) + (\bar{p} - t) \ln \left( \frac{\bar{p}}{\bar{p}-t} \right)$
- Observe that $f(0) = 0$
- Note that

$$f'(t) = \ln \left( \frac{p}{p+t} \right) - 1 - \ln \left( \frac{\bar{p}}{\bar{p}-t} \right) + 1$$

$$= \ln \left( \frac{p}{p+t} \right) - \ln \left( \frac{\bar{p}}{\bar{p}-t} \right)$$

- Observe that $f'(0) = 0$
- Note that

$$f''(t) = - \frac{1}{p+t} - \frac{1}{\bar{p}-t} = - \frac{1}{(p+t)(\bar{p}-t)}$$
Now, we use Taylor Series expansion around $x_0 = 0$. There exists a positive constant $c \in (0, t]$ such that:

$$f(t) = f(0) + f'(0) \cdot t + f''(c) \cdot \frac{t^2}{2}$$

$$= -\frac{1}{(p + c)(\bar{p} - c)} \cdot \frac{t^2}{2}$$

$$\leq -2t^2$$

For the last step, we used AM-GM inequality:

$$\sqrt{(p + c)(\bar{p} - c)} \leq \frac{(p + c) + (\bar{p} - c)}{2} = \frac{1}{2}$$

Recall that we had defined $f(t) = \ln E$, so $f(t) \leq -2t^2$ implies $E \leq \exp(-2t^2)$. This completes the proof of the last part of Chernoff Bound.