CS 70 Discrete Mathematics and Probability Theory
Fall 2015 Jean Walrand 13B Discussion Solutions

Chebyshev's Inequality $P[|X - E[X]| > a] \leq \frac{Var(X)}{a^2}$ for all a > 0.

Chernoff bound: $P[X \geq \alpha n] \leq e^{-n\Phi(\alpha)}$ where X is a summation of independent random variables X_i for $i = 1 \ldots n$ and $P[X_i = 1] = p$ and $\Phi(\alpha) = \alpha \ln(\frac{\alpha}{p}) + (1-\alpha)\ln(\frac{1-\alpha}{1-p})$

1. Homework Polling

   Suppose Prof. Walrand wants to poll the CS 70 students about whether the homeworks have been too hard recently. Suppose everyone is comfortable enough to answer honestly, either no (0) or yes (1). Let the true fraction of students who think the homework is too hard be p. Let the response of the i^{th} polled student be X_i. Prof. Walrand would like to poll n students, with n large enough that the probability of estimating p to within $\pm 2\%$ is at least 90%.

   (a) Let $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the average of the n responses. What is the expectation of M_n? What is the event that we are interested in whose probability we would like to be at least 90%? Draw a picture of the distribution of $M_n - p$ and mark the region that corresponds to the event of interest.

   Answer:

$$X_i = \begin{cases} 1 & \text{hw is too hard} \\ 0 & \text{hw is not too hard} \end{cases}$$

   X_i is i.i.d Bernoulli r.v. with $\Pr[X_i = 1] = p$. So, $M_n \sim \frac{1}{n}Binom(n, p)$:

$$E[M_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\cdot(np) = p$$

   We want

$$\Pr[|M_n - p| < 0.02] > 0.9$$

   Or equivalently,

$$\Pr[|M_n - p| \geq 0.02] \leq 0.1$$

(b) Now use Chebyshev's inequality to find a safe n regardless of what p is.

Answer:

$$Var[M_n] = \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i] = \frac{p(1-p)}{n}$$

$$\Pr[|M_n - p|] \geq 0.02] \leq \frac{Var[M_n]}{0.02^2} = \frac{p(1-p)}{(0.02)^2 n} \leq 0.1$$

$$n \geq 25000p(1-p) \leq 6250$$

Here, $p(1-p) \leq 1/4$ when equality holds of $p = 1/2$.

(c) What if instead of wanting an accuracy of $\pm 2\%$ we wanted a relative error of 2%. This means that if the true value was p, we want the answer we get to be within $[0.98p, 1.02p]$. Can we pull that off using Chebyshev's inequality? Do you think we could pull that off with a single universal choice of n that does not depend on (the unknown) p?

Answer: We want

$$\Pr[0.98p < M_n < 1.02p] > 0.9$$

$$\Pr[|M_n - p| \geq 0.02p] \leq 0.1$$

So, plugging in the $Var[M_n]$ in the Chebyshev's inequality, we have

$$\Pr[|M_n - p| \geq 0.02p] \leq \frac{Var[M_n]}{(0.02p)^2} = \frac{1-p}{0.02^2 np} \leq 0.1$$

Solving this, we get

$$n \geq 25000\frac{1-p}{p}$$

It depends on p as there is no simple bounds on $\frac{1-p}{p} = 1/p - 1$.

2. Confidence

You have a random variable X whose expectation you'd like to estimate. You have the ability to draw samples from X's distribution, and you know nothing about $\mathrm{E}(X)$, and you only know that $\mathrm{Var}(X) \leq 10$. (You may assume $\mathrm{Var}(X) = 10$, since this is the worst case.) To estimate $\mathrm{E}(X)$, you take n i.i.d. samples of X and average them. Using the bound proved in lecture based on Chebyshev's Inequality (equation (3) in lecture note 17):

(a) Suppose you take 1000 samples. How confident are you that your estimate is within an absolute error of 0.5?

Answer: We have $n = 1000$ samples, and $\varepsilon = 0.5$ error bound. We have $n \geq \frac{\sigma^2}{\varepsilon^2 \delta}$, and wish to find a bound on δ, our confidence. Since $n > 0$ and $\delta > 0$, we manipulate this equation and see that $\delta \geq \frac{\sigma^2}{\varepsilon^2 n} = \frac{\sqrt{10}^2}{0.5^2 \cdot 1000} = 0.04$. This means we are 96% sure in our answer.

(b) Suppose instead that you want an absolute error of at most 2 and a confidence parameter of 0.02 (you want to be "98% confident"). How many samples do you need?

Answer: We have $\varepsilon = 2$ error bound, and $\delta = 0.02$ confidence. We have $n \geq \frac{\sigma^2}{\varepsilon^2 \delta}$ and wish to find a bound on n, our number of samples. Plugging in our numbers, $n \geq \frac{\sqrt{10}^2}{2^2 \cdot 0.02} = 125$, so we should take at least 125 samples to achieve our desired error and confidence.

(c) Suppose instead that you take 2500 samples and you want a confidence parameter of 0.1 ("90% confident"). What absolute error bound will you get with this confidence?

Answer: We have $n = 2500$ samples and want $\delta = 0.1$ confidence. We have $n \geq \frac{\sigma^2}{\varepsilon^2 \delta}$, and wish to find a bound on ε, our error bound. Since $n > 0$ and $\varepsilon > 0$, we manipulate this equation and see that $\varepsilon \geq \sqrt{\frac{\sigma^2}{\delta \cdot n}} = \sqrt{\frac{\sqrt{10}^2}{0.1 \cdot 2500}} = 0.2$, so our error bound is 0.2.

3. Chernoff vs. Chebyshev

Here we want to compare Chernoff's bound and the bound you can get from Chebyshev's inequality.

(a) Consider the experiment of flipping a fair coin 100 times. Let S be the number of heads you obtain in your experiment. The exact probability of $S \geq 80$ is $5.5795 \cdot 10^{-10}$. What estimate does Chebyshev's bound give for the probability of seeing at least 80 heads in 100 coin flips? What about Chernoff's bound? Compare your answers and see which one is closer to the actual value.

Answer: From Chebyshev's inequality:

$$Pr(S \geq 80) = Pr(S - 50 \geq 30) \leq Pr(|S - 50| \geq 30) \leq \frac{\mathrm{Var}(S)}{30^2} = \frac{100 \cdot 0.5(1 - 0.5)}{30^2} = 0.0278$$

.

For Chernoff's bound:
We have a similar case here as in the class notes (Note 19, extended example). Our coin flips are i.i.d. Bernoulli(0.5). We can consider S_i to be the indicator random variable that is 1 if flip i is

heads, and 0 if it is tails. Then $S = S_1 + S_2 + \ldots + S_n$, and noting that $P(S \geq 80) = P(\frac{1}{100} \cdot S \geq 0.8)$, we can use the Chernoff bound:

$$P(S \geq 80) = P\left(\frac{1}{100}S \geq 0.8\right) \leq e^{-100\Phi_{X_1}(0.8)}.$$

Recall that since our problem here is the same as the notes example, we can use the Kullback-Liebler Divergence:

$$\Phi_{X_1}(\alpha) = D(\alpha||p) = \alpha\ln(\alpha/p) + (1-\alpha)\ln((1-\alpha)/(1-p)),$$

so we get the bound

$$P(S \geq 80) \leq e^{-100D(0.8||0.5)} = e^{-100(0.8\ln(0.8/0.5)+0.2\ln((0.2)/(0.5)))} = 4.2580 \cdot 10^{-9}.$$

This is a much better bound than what we got from Chebyshev's inequality.

(b) Now back to the setting with general n and α, write down Chernoff's bound in the form c^n, where c is an expression that only contains α and not n.

Answer:

Let's try to get our Chernoff bound in the form c^n.

$$\begin{aligned}
e^{-n\left(\alpha\ln\left(\frac{\alpha}{p}\right)+(1-\alpha)\ln\left(\frac{1-\alpha}{1-p}\right)\right)} &= e^{n\left(\ln\left(\left(\frac{\alpha}{p}\right)^{-\alpha}\right)+\ln\left(\left(\frac{1-\alpha}{1-p}\right)^{\alpha-1}\right)\right)} \\
&= \left(e^{\ln\left(\left(\frac{\alpha}{p}\right)^{-\alpha}\right)+\ln\left(\left(\frac{1-\alpha}{1-p}\right)^{\alpha-1}\right)}\right)^n \\
&= \left(e^{\ln\left(\left(\frac{\alpha}{p}\right)^{-\alpha}\right)}e^{\ln\left(\left(\frac{1-\alpha}{1-p}\right)^{\alpha-1}\right)}\right)^n \\
&= \left(\left(\frac{\alpha}{p}\right)^{-\alpha}\left(\frac{1-\alpha}{1-p}\right)^{\alpha-1}\right)^n
\end{aligned}$$

So, we have $c = \left(\left(\frac{\alpha}{p}\right)^{-\alpha}\left(\frac{1-\alpha}{1-p}\right)^{\alpha-1}\right)$.

This shows that for a fixed value of α and p, Chernoff's bound decays exponentially in n.

(c) Show that Chebyshev's bound on $\Pr[|S-\mu| \geq \alpha\mu]$ is also a bound on $\Pr[S \geq (1+\alpha)\mu]$.

Answer:

$P(|S-\mu| \geq \alpha\mu) = P(S-\mu \geq \alpha\mu) + P(\mu - S \geq \alpha\mu)$

$$\geq P(S-\mu \geq \alpha\mu)$$

$$= P(S \geq (1+\alpha)\mu)$$

(d) Using the previous part above, write a bound on $\Pr[S \geq (1+\alpha)\mu]$ of the form γn^{β}, where γ and β are numbers that do not depend on n. What does this tell you about Chebyshev's inequality vs. Chernoff's inequality?

Answer: From Chebyshev's inequality, we have:

$$P(S \geq (1+\alpha)\mu) \leq P(|S-\mu| \geq \alpha\mu) \leq \frac{np(1-p)}{(\alpha np)^2} = \frac{1-p}{p\alpha^2} \cdot n^{-1}$$

,
where $\gamma = \dfrac{(1-p)}{p\alpha^2}$ and $\beta = -1$.

This shows that Chebyshev's inequality decays like n^{β}. In general an exponential decay (which you get from Chernoff's) is much faster than a polynomial decay (the one you get from Chebyshev's).