Lecture - 25

Friday, 7 October 2016 (16:15- 17:05)

Chernoff Bound

1 Need for Chernoff Bound

Assume Dr. Sudarshan is playing Chess with i of his students. He either wins or loses a game. We model this situation with the help of random variables as follows :
Let X_i be a random variable associated with the student i.

$$X_i = \begin{cases} 1 & \text{if Dr. Sudarshan wins the game with } i^{th} \text{ student} \\ 0 & \text{otherwise} \end{cases}$$

The total number of games won by Dr. Sudarshan can now be expressed as a random variable X, where $X = X_1 + X_2 + ... + X_n$

Expected number of games won by him = $E[X] = E[X_1] + E[X_2] + ... + E[X_n]$ (from Linearity of Expectations)

If his probability of winning with the i^{th} student is p_i, then

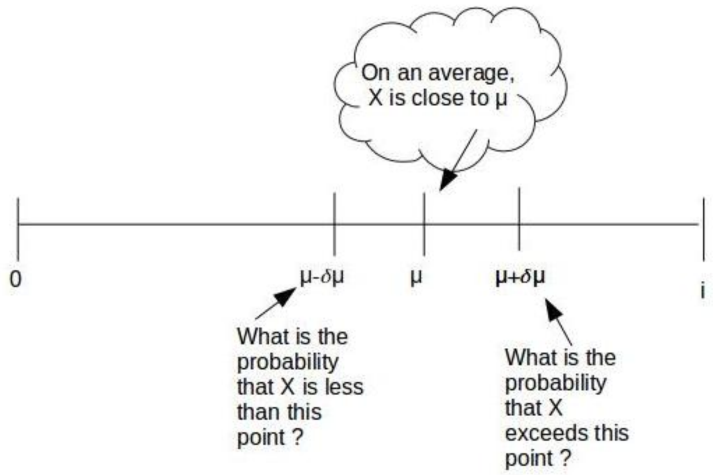
$E[X] = p_1 + p_2 + ... + p_n$ (Since, $E[X_i] = p_i$, as it is an indicator random variable)

We now know $E[X]$. We also know that on an average, the value of X for an experiment will be close to $E[X]$, but how much close? How much the value of X can deviate from its expected value. Let $E[X] = \mu$, then we are interested in finding the following

$$\boxed{Pr(X \geq \mu + \mu\delta) = ?}$$

$$\boxed{Pr(X \leq \mu - \mu\delta) = ?}$$

We derive the first expression in this lecture. The proof for the second inequality is left as a homework exercise.



2 Some Prerequisites for the Derivation

2.1 Markov's Inequality

If there is a random variable Y taking values from the set $\{1, 2, ..., n\}$, then

$$\boxed{Pr(Y \geq a) \leq \frac{E[Y]}{a}}$$

The proof is straightforward and is left as an exercise problem.

2.2 Moment Generating Function

If Y is a random variable taking values from the set {1, 2, ..., n} then, e^{tY} is the corresponding generating function for Y.

$$\boxed{Pr(Y \geq a) = Pr(e^{tY} \geq e^{ta})}$$

2.3 $1 + y \leq e^y$ if y is positive

$e^y = 1 + y + \frac{y^2}{2!} + +\frac{y^3}{3!} + ...$

Hence,

$$1 + y \leq e^y$$

2.4 $E[e^{tY}] = p(e^t - 1) + 1$ for an indicator random variable Y which is 1 with probability p

Let $Y = 1$ with probability p and $Y = 0$ with probability $1 - p$.

If $Y = 1$, then $e^{tY} = e^t$. If $Y = 0$, then $e^{tY} = 1$.

So, $e^{tY} = e^t$ with probability p and $e^{tY} = 1$ with probability $1 - p$.

Hence, $E[e^{tY}] = p(e^t) + (1 - p)$

$$\boxed{E[e^{tY}] = p(e^t - 1) + 1}$$

2.5 $E[Y \times Z]$ where Y and Z are independent random variables

$$\boxed{E[Y \times Z] = E[Y] \times E[Z]}$$

3 Chernoff Bound- Inequality and Derivation

3.1 Inequality

As explained in section 1, let X be a random variable which is the sum of n independent random variables X_1, X_2, ..., X_n.

$X_i = 1$ with probability p_i.

Let $E[X] = \mu$, then

$$\boxed{pr(X \geq (1+\delta)\mu) \leq \frac{e^{\delta\mu}}{(1+\delta)^(1+\delta)\mu}}$$

3.2 Derivation

$pr(X \geq (1+\delta)\mu)$

$= pr(e^{tX} \geq e^{t(1+\delta)\mu})$ (From subsection 2.2)

$\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}}$ (From subsection 2.1)

$= \frac{E[e^{t\sum_{i=1}^{n}(X_i)}]}{e^{t(1+\delta)\mu}}$ (Since, $X = X_1 + X_2 + ... + X_n$)

$= \frac{E[\prod_{i=1}^{n} e^{t(X_i)}]}{e^{t(1+\delta)\mu}}$ (Since, $e^{a+b+c} = e^a e^b e^c$)

$= \frac{\prod_{i=1}^{n} E[e^{t(X_i)}]}{e^{t(1+\delta)\mu}}$ (From subsection 2.5)

$= \frac{\prod_{i=1}^{n}(1+p_i(e^t-1))}{e^{t(1+\delta)\mu}}$ (From subsection 2.4)

$= \frac{\prod_{i=1}^{n} e^{(p_i(e^t-1))}}{e^{t(1+\delta)\mu}}$ (From subsection 2.3)

$= \frac{e^{\sum_{i=1}^{n}(p_i(e^t-1))}}{e^{t(1+\delta)\mu}}$ (Since, $e^{a+b+c} = e^a e^b e^c$)

$= \frac{e^{(e^t-1)\sum_{i=1}^{n}(p_i)}}{e^{t(1+\delta)\mu}}$ (Since, $e^{a+b+c} = e^a e^b e^c$)

$= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}$ (Since, $E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i$)

Put $t = \log(1+\delta)$, Intuition for choosing this value of t is- To maximise this function, differentiate w.r.t. t and equate to 0.

$pr(X \geq (1+\delta)\mu) \leq \frac{e^{(e^{\log(1+\delta)}-1)\mu}}{e^{\log(1+\delta)\times(1+\delta)\mu}}$

$= \frac{e^{(1+\delta-1)\mu}}{e^{\log(1+\delta)^{(1+\delta)\mu}}}$

$= \frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}}$

Hence,

$$\boxed{pr(X \geq (1+\delta)\mu) \leq \frac{e^{\delta\mu}}{(1+\delta)^(1+\delta)\mu}}$$

A Moment Generating Function

If Y is a random variable taking values from the set {1, 2, ..., n} then, e^{tY} is the corresponding generating function for Y.

$E[Y] = \sum_{i=1}^{n} i \times pr(Y = i)$

$E[e^{tY}] = E[tY + \frac{(tY)^2}{2!} + \frac{(tY)^3}{3!} + \frac{(tY)^4}{4!} + ... + \frac{(tY)^n}{n!}]$

$= tE[Y] + \frac{t^2}{2!}E[Y^2] + \frac{t^3}{3!}E[Y^3] + ...\frac{t^n}{n!}E[Y^n]$

The i^{th} term in the above equation is called the i^{th} moment.

The reason why the function e^{tY} is called the moment generating function is - We can get different moments of Y. We differentiate the equation i times and equate it to zero. This gives us the expected value i.e. the i^{th} moment.