Math

Random variables and concentration

Linearity of expectation and variance:

> For R.V X, Y:
>
> Expectation:
>
> > $\mathbb{E}[X] = \sum_{s \in S} Pr[X = s] * S$
> > $\mathbb{E}[\alpha X] = \alpha \mathbb{E}[X]$
> > $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
>
> Variance：
>
> > $\mathbb{V}\text{ar}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$
> > $\mathbb{V}\text{ar}[\alpha X] = \alpha^2 \mathbb{V}\text{ar}[X]$
> > $\mathbb{V}\text{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
>
> When X and Y are independent:
>
> > $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$
> > $\mathbb{V}\text{ar}[X + Y] = \mathbb{V}\text{ar}[X] + \mathbb{V}\text{ar}[Y]$
>
> Random event A,B:
>
> > $P(A \cup B) = P(A + B) = P(A) + P(B) - P(AB)$
> > $P(A \cap B) = P(AB) = P(A, B) = P(A|B)P(B)$
> >
> > One trick to identify if A and B are independent:
> >
> > > if we know B happens or not, does we have any indication that A will happen or not
> >
> > When A and B are independent:
> >
> > > $P(A|B) = P(A)$
> > > $P(A \cap B) = P(A) * P(B)$

Indicator random variables and how to use them

> when to use:
> know $\mathbb{E}[d_{i,j}]$, D contains a lot of events and they are dependent
> $d_{i,j} = 0/1 \quad D = \sum_{i,j} d_{i,j}$

Inequality

> Union bound:
>
> > $P[A_1 \cup A_2 \cup A_3 \cup ... \cup A_k] \leq P[A_1] + P[A_2] + ... + Pr[A_k]$
>
> How well does a $R.V.\ X$ concentrate around it's expectation $\mathbb{E}[X]$:
>
> Markov's Inequality:

> For any random variable X which only takes non-negative values any positive t:
> $$Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t} \Leftrightarrow Pr[X] \geq \alpha\mathbb{E}[X]] \leq \frac{1}{\alpha}$$
>
> Prove:
> $E[X] = \sum_{x=0}^{\infty} xPr[X=x] \geq \sum_{x=t}^{\infty} xPr[X=x] \geq \sum_{x=t}^{\infty} tPr[X=x] = t \cdot \sum_{x=0}^{\infty} Pr[X=x] = t \cdot Pr[X \geq t]$

Chebyshev's Inequality

> r.v. X with expectation $\mathbb{E}[x]$ and variance $\sigma^2 = \mathbb{V}\mathrm{ar}[X]$. Then for any k > 0:
> $$Pr[|X - \mathbb{E}[X] \geq k\sigma|] \leq \frac{1}{k^2}$$
>
> Ad: X could be anything, two side
> Dis: need to know variance
>
> using linearity of variance
> variance deduction: i.i.d, means of many trails

Gaussian Tail Bound

> For $X \sim N(\mu, \sigma^2)$: $\quad Pr[|X - \mathbb{E}X| \geq \alpha\sigma] \leq O(e^{-\alpha^2/2})$

Theorem(CLT-Informal)

> Any sum of independent, identically distributed r.v.'s $X_1, \dots X_k$ with μ and finite variance σ^2 converges to a Gaussian r.v. with mean $k\mu$ and variance $d\sigma^2$ as $k \to \infty$:
> $$S = \sum_{i=1}^{n} X_i \to N(k\mu, k\sigma^2)$$

Chernoff Bound

> Let $X_1, X_2, \dots X_k$ be independent {0,1} - valued R,V. and let $p_i = \mathbb{E}[X_i]$, where $0 < p_i < 1$. Then the sum $S = \sum_{i=1}^{k} X_i$, which has mean $\mu = \sum_{i=1}^{k} p_i = \mathbb{E}[S]$ satisfies :
>
> $$Pr[S \geq (1+\epsilon)\mu] \leq e^{\frac{=\epsilon^2\mu}{2+\epsilon}}$$
>
> and for $0 < \epsilon < 1$
> $$Pr[S \leq (1-\epsilon)\mu] \leq e^{\frac{\epsilon^2\mu}{2}}$$
> Corollary of Chernoff bound:
> for $0 < \Delta < 1$, $\quad Pr[|S - \mu| \geq \Delta\mu] \leq 2e^{-\Delta\mu/3}$

Bernstein Inequality

> Let $X_1, X_2, \dots, X_k$ be independent random variables with each $X_i \in [-1, 1]$. Let $\mu_i = \mathbb{E}[X_i]$ and $\sigma_i^2 = \mathbb{V}\mathrm{ar}[X_i]$. Let $\mu = \sum_i \mu_i$ and $\sigma^2 = \sum_i \sigma_k^2$. Then, for $\alpha \leq \frac{1}{2}\sigma$, $S = \sum_i X_i$ satisfies:
> $$Pr[|S - \mu| > \alpha\sigma] \leq exp(-\frac{\alpha^2}{4})$$

Hoeffding Inequality

> Let $X_1, X_2, \dots, X_k$ be independent random variables with each $X_i \in [a_i, b_i]$. Let $\mu_i = \mathbb{E}[X_i]$ and $\sigma_i^2 = \mathbb{V}\mathrm{ar}[X_i]$. Let $\mu = \sum_i \mu_i$ and $\sigma^2 = \sum_i \sigma_k^2$. Then, for any $\alpha > 0$, $S = \sum_i X_i$ satisfies:
> $$Pr[|S - \mu| > \alpha] \leq 2\ exp(-\frac{\alpha^2}{\sum_{i=1}^{k}(b_i - a_i)^2})$$

Variance is a natural measure of central tendency
q^{th} central moment: $\mathbb{E}[(X - \mathbb{E}X)^q]$

Some numbers

$log_2 1billion = log_2 1000000 \approx 20$

0! = 1