Lecture # 10

Chapter 5: Probability Bounds

1 The Union Bound and its Extensions

a. Union Bound

ecall for two events A and B

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

$$\leq P[A] + P[B]$$

$\rightarrow$ Generalization to multiple events

Union Bound (known also as the Boole's inequality)

For any events $A_1, A_2, \dots, A_n$, we have

$$P[A_1 \cup A_2 \dots \cup A_n] \leq P[A_1] + P[A_2] + \dots + P[A_n]$$

→ It is widely in the area of random graphs

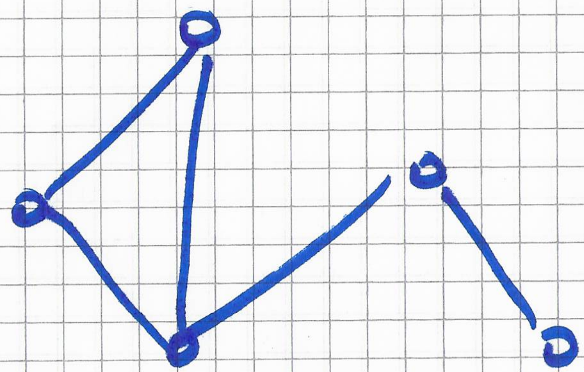
Example: Consider an Erdös-Rényi Graph model $G(n, p)$

where:

- n: number of nodes in the graph
- Each pair of nodes are connected by an edge

with probability p.

→ Note that occurence of edges are independent from each other.



$n=5$

Q: what is the probability that there exists an isolated node in the graph (i.e a node not connected to any other node)

A: ~~Let~~ Consider $G(n,p)$ model

Let B_n: Event that $G(n,p)$ has at least one isolated node

Let A_i be the event that the i-th node is isolated

$$B_n = \bigcup_{i=1}^{n} A_i$$

$\Downarrow$ Union Bound

$$P[B_n] = P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i]$$

By symmetry, we can easily see that

$$P[A_i] = P[A_j] = P(A_1)$$

Thus

$$\mathbb{P}[B_n] \leq n\, \mathbb{P}[A_1]$$

Given that the connections are independent

$$\mathbb{P}[A_1] = (1-p)^{n-1}$$

Therefore

$$\boxed{\mathbb{P}[B_n] \leq n\,(1-p)^{n-1}}$$

b_ Extensions

We can extend the union bounds and obtain the lower and upper bound on probability of union of events

. Generalization of Union Bound: Bonferroni Inequalities.

→ For any events $A_1, A_2, \dots, A_n$, we have:

. $$P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i]$$

. $$P\left[\bigcup_{i=1}^{n} A_i\right] \geq \sum_{i=1}^{n} P[A_i] - \sum_{i<j} P[A_i \cap A_j]$$

. $$P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i] - \sum_{i<j} P[A_i \cap A_j] + \sum_{i<j<k} P[A_i \cap A_j \cap A_k]$$

In general, if you write an odd number of terms then you get an upper bound and if you write an even number of terms you get a lower bound.

Proof: $\to$ Hwk #6 (Pb #1)

Hint: start writing the inclusion-exclusion formula. If you stop at the first term, you obtain an upper-bound on the union probability. If you stop at the second term, you obtain a lower bound. If you stop at third term you get again a refined upper-bound, ect...

(c) Expected Value of the number of Events

→ The union bound formula is also equal to the expected value of the number of occured events.

→ Let $A_1, A_2, \ldots, A_n$ be any events.

Define the indicator random variables

$X_1, X_2, \ldots, X_n$ as

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

→ If we define $X = X_1 + X_2 + \cdots + X_n$ then

X shows the number of events that actually occurs

$$E[X] = E[X_1 + X_2 + \cdots + X_n]$$

$$= E[X_1] + E[X_2] + \cdots + E[X_n]$$

$$= P[A_1] + P[A_2] + \cdots + P[A_n]$$

which is indeed the right-hand-side of the union bound.

→ So with reference to previous Example, we can conclude that the expected number of isolated nodes in a graph randomly generated according to $G(n,p)$ is also equal to

$$\boxed{E[X] = n(1-p)^{n-1}}$$

5.2 Markov, Cheybeshev, and Chernoff Inequalities

a- Markov Bound

Th: If X is a non-negative RV, then

$$P[X \geq a] \leq \frac{E[X]}{a} \quad \text{for any } a > 0$$

Remark: This inequality is useful only when $a > E[X]$

Proof: Since X is a non-negative RV

$$f_X(x) = 0 \quad \text{for all } x < 0$$

$$E[X] = \int_{-\infty}^{+\infty} x \, f_X(x) \, dx = \int_{0}^{+\infty} x \, f_X(x) \, dx$$

$$= \underbrace{\int_0^a \underbrace{\underset{\geq 0}{x}\, \underset{\geq 0}{f_X(x)}}_{\geq 0} \, dx}_{\geq 0} + \int_a^{+\infty} x \, f_X(x) \, dx$$

$$\Rightarrow \quad E[X] \geq \int_a^{+\infty} \underset{x \geq a}{\underset{\downarrow}{x}} \, f_X(x) \, dx$$

$$\geq \int_a^{+\infty} a \, f_X(x) \, dx$$

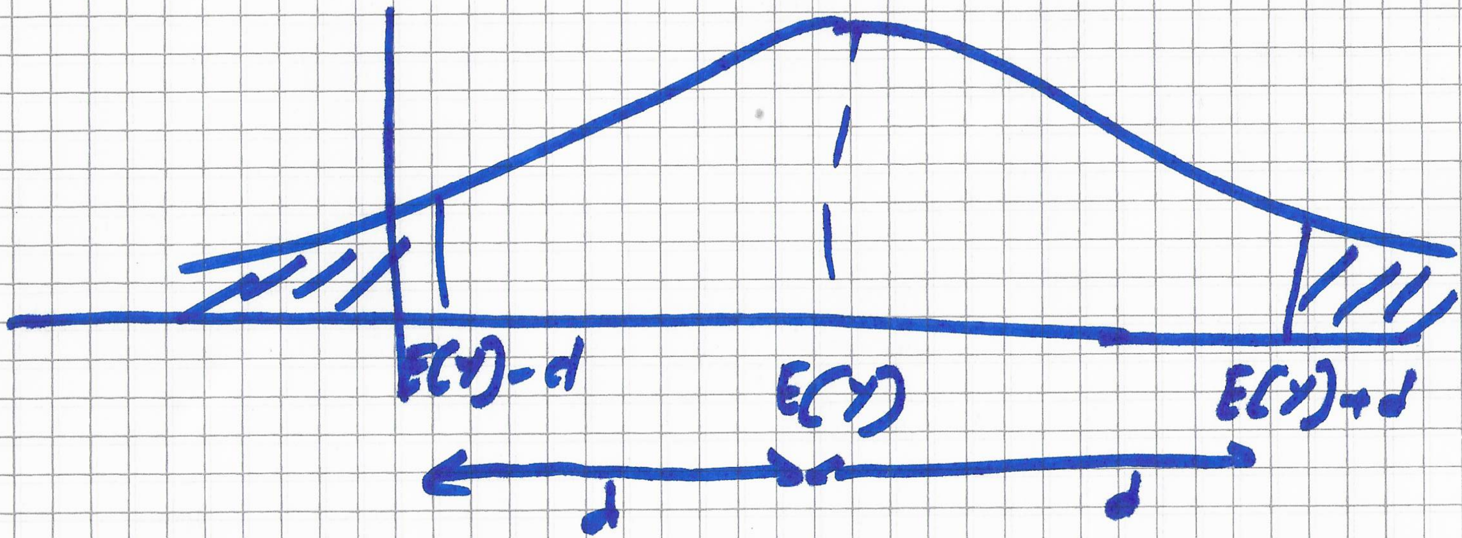
$$= a \int_a^{+\infty} f_X(x) = a \, Prob[X \geq a]$$

$$\Leftrightarrow \quad \boxed{Prob[X \geq a] \leq \frac{E[X]}{a} \quad \text{for } a > 0}$$

b_ Cheybeshev Inequality

Th: For any arbitrary RV Y and a constant d we have

$$\text{Prob}\left[\,|Y - E[Y]| \geq d\,\right] \leq \frac{Var[Y]}{d^2}$$



Proof:

Let $X = (Y - E[Y])^2 \geq 0$

Apply Markov Inequality to X

$$Prob\left[\,|Y - E[Y]|^2 \geq a\,\right] \leq \frac{\overbrace{E[(Y - E[Y])^2]}^{=\,Var(Y)}}{a}$$

Let $a = d^2$

$$Prob\left[\,|Y - E[Y]| \geq d\,\right] \leq \frac{Var[Y]}{d^2}$$

QED

C- Chernoff Inequality

For an arbitrary RV X and a constant c

$$\mathrm{Prob}[X \geq c] \leq \min_{s \geq 0} \left\{ e^{-sc} \mathcal{M}_X(s) \right\}$$

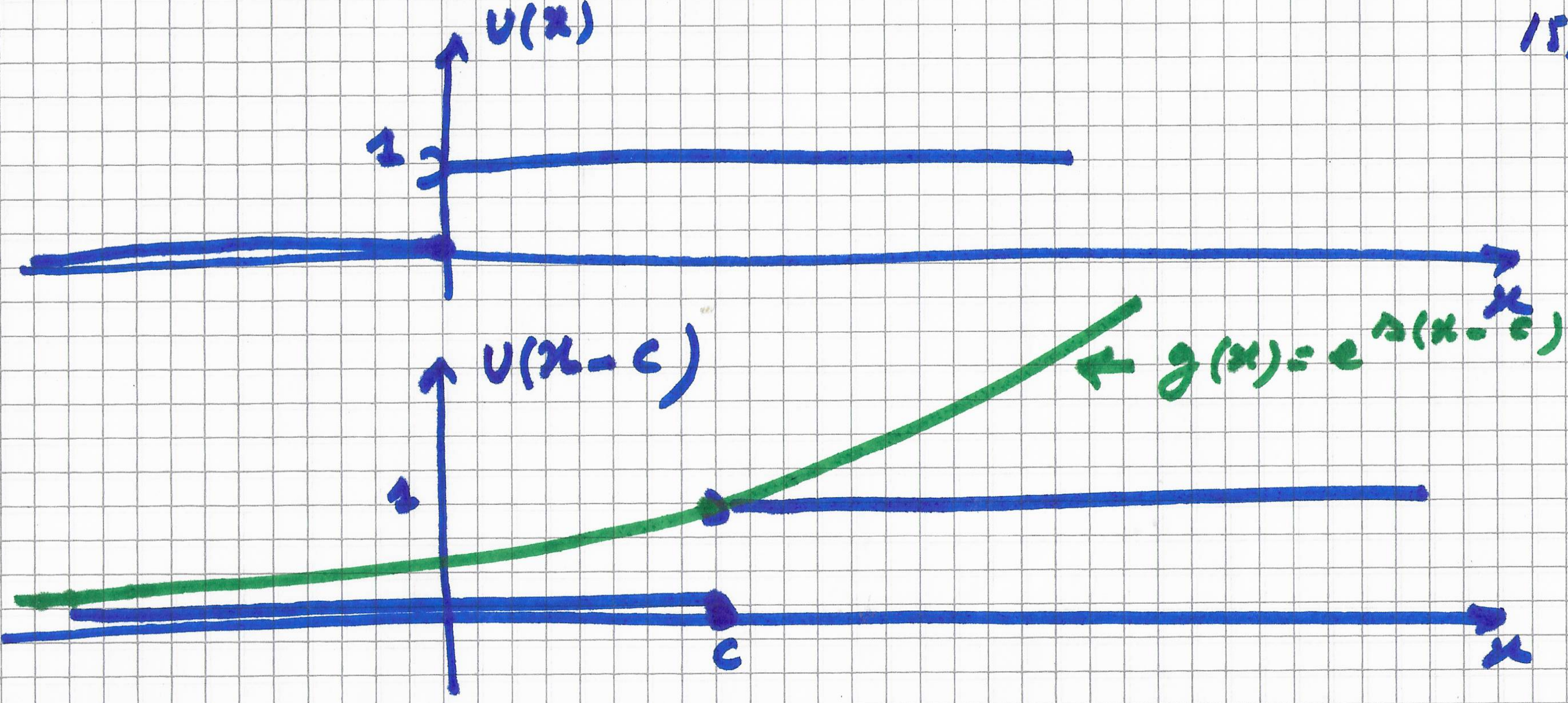
where $\mathcal{M}_X(s) = E[e^{sX}]$ is the MGF of RV X.

Proof:

$$\mathrm{Prob}[X \geq c] = \int_c^{+\infty} f_X(x)\, dx$$

$$= \int_{-\infty}^{+\infty} f_X(x)\, u(x - c]\, dx$$

shifted unit step function



For all x $\quad U(x-c) \leq \boxed{e^{s(x-c)}}$

$s \geq 0$

Thus $P[X \geq c] = \int_{-\infty}^{+\infty} f_X(x)\, U(x-c)\, dx$

$$\leq \int_{-\infty}^{+\infty} f_X(x)\, e^{s(x-c)}\, dx$$

$$= e^{-sc} \underbrace{\int_{-\infty}^{+\infty} f_X(x) \cdot e^{sx} \, dx}_{E[e^{sx}] = M_X(s)}$$

Therefore

$$\text{Prob}[X \geq c] \leq e^{-sc} M_X(s)$$

and this inequality is valid for any $s \geq 0$

So in particular, it holds for s^* for which

$$e^{-s^* x} M_X(s^*) \text{ is minimum}$$

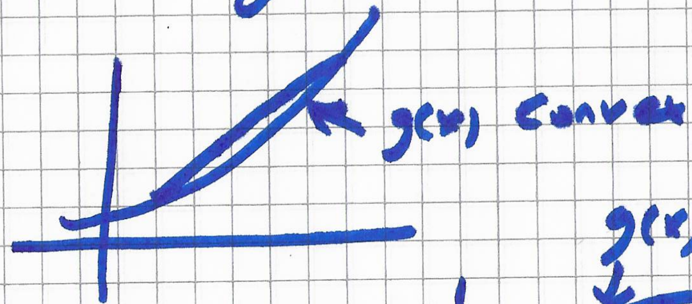
QED

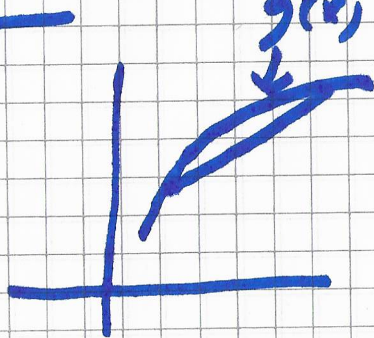
5.3 Jensen's Inequality

Th: If $g(x)$ is a convex function on R_x and $E[g(x)]$ and $g(E[x])$ are finite, then

$$E[g(x)] \geq g[E[x]]$$

Recall: A twice differentiable function $g(x)$ is convex iff $g''(x) \geq 0$





Remark: If $g(x)$ is concave $(g''(x) \leq 0)$

$\Downarrow$

$$E[g(x)] \leq g(E[x])$$

Example

Let X be a positive RV

Compare $E[X^a]$ with $(E[X])^a$ for $a \in \mathbb{R}$

Note . $a=0$ $\quad E[X^0] = 1 = (E[X])^0 = 1$

. $a=1$ $\quad E[X^1] = E[X] = (E[X])^1 = E[X]$

Let us assume $a \neq 0$ and $a \neq 1$

If $g(x) = x^a$, $\quad g'(x) = a\, x^{a-1}$,

$$g''(x) = a\,(a-1)\, x^{a-2} \quad \text{with } x \geq 0$$

If $\underline{a<0}$ $\quad g''(x) \geq 0 \implies g(x)$ is convex

If $\underline{a>1}$

If $0 < a < 1 \Rightarrow g''(x) \leq 0 \Rightarrow g(x)$ is concave

with these results and applying Jensen's inequality

- for $a < 0$ and $a > 1$

$$E[X^a] \geq (E[X])^a$$

- for $0 < a < 1$

$$E[X^a] \leq (E[X])^a$$

5.4 Hölder, Cauchy-Shwartz, Minkowski, and Lyapunov Inequalities

a- Hölder Inequality

Th: If $p, q > 1$ and $p^{-1} + q^{-1} = 1$ then

$$E[|XY|] \le \left(E[|X|^p]\right)^{1/p} \left(E[|Y|^q]\right)^{1/q}$$

Proof: Hwk # 6.

b- Cauchy-Shwartz Inequality

If $p = q = 2$ in Hölder Inequality

$$E[\,|XY|\,] \le \sqrt{E[X^2]} \cdot \sqrt{E[Y^2]}$$

$$E[\,|XY|\,] \le \sqrt{E[X^2]\,E[Y^2]} \iff$$

$$\left(E[XY]\right)^2 \le E[X^2]\,E[Y^2]$$

c. Minkowski Inequality

Th: $$\left(E[\,|X+Y|^p\,]\right)^{1/p} \le \left(E[\,|X|^p\,]\right)^{1/p} + E[\,|Y|^p\,]^{1/p}$$

Proof: $Z = |X+Y|$

$$E[Z^p] = E[\underbrace{Z}_{|X+Y|} \cdot Z^{p-1}]$$

$$\leq E[|X|\, Z^{p-1}] + E[|Y|\, Z^{p-1}]$$

Hölder

$$\leq \left(E[|X|^p]\right)^{1/p} \left(E[Z^{(p-1)\cdot q}]\right)^{1/q} +$$

$$\left(E[|Y|^p]\right)^{1/p} \cdot \left(E[Z^{(p-1)\cdot q}]\right)^{1/q}$$

but $\frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow \frac{q+p}{qp} = 1 \Leftrightarrow q + p = qp$

$$pq - q = p$$

Thus

$$E[Z^p] \leq E[\,|X|^p]^{1/p} \cdot \left(E[Z^p]\right)^{1/q} +$$

$$E[\,|Y|^p]^{1/p} \left(E[Z^p]\right)^{1/q}$$

Divide both sides with $\left(E[Z^p]\right)^{1/q}$ leads to

$$\frac{E[Z^p]}{\left(E[Z^p]\right)^{1/q}} \leq \left(E[\,|X|^p]\right)^{1/p} + \left(E[\,|Y|^p]\right)^{1/p}$$

$$\parallel$$

$$\left(E[Z^p]\right)^{1-\frac{1}{q}} \leq \quad " \quad\quad\quad "$$

but $\frac{1}{p}+\frac{1}{q}=1 \Rightarrow 1-\frac{1}{q}=\frac{1}{p} \Rightarrow$

$$\Big(E\big[\underset{\substack{\shortparallel \\ |X+Y|}}{Z}^p\big]\Big)^{1/p} \leq \big(E[|X|^p]\big)^{1/p} + \big(E[|Y|^p]\big)^{1/p}$$

QED.

d. Lyapunov's Inequality

$$\big(E[|Z^r|]\big)^{1/r} \geq \big(E[|Z^s|]\big)^{1/s}$$

for $r \geq s \geq 0$

Proof: Hwk #6

Implication: If RV Z has a finite r^{th} moment then all s^{th} moments are finite also for positive $s \leq r$.