CAS CS 365
Lecture 4

Probability
1. Markov's inequality
    a. $$\boxed{\Pr(X \geq t) \leq \frac{\mathbb{E}[X]}{t}}$$
    b. The inequality is tight → Cannot be improved
    c. Example)
        i. Let X be a non-negative RV. Then, $\boxed{\Pr(X \geq k\mathbb{E}[X]) \leq \frac{1}{k}}$
        ii. For example X~Bin(1000,0.1).
        iii. X = # heads of based on the coin where p = 0.1
        iv. Then E[x] = np = 1000 * 0.1 = 100. → the expected value is non negative → can apply markov's inequality
        v. Pr(x=k) = nCk * (1-p)^(1-k) * (p)^k = nCk * (0.9)^(1-k) * (p)^k
        vi. Markov's inequality: $\Pr(X \geq 400) \leq \frac{1}{4}$
        vii. The bound is not tight, we can get much tighter bounds → can make the upper bound smaller
2. Chebyshev's inequality
    a. Let μ = E[x], o^2 = Var[x]. Then $\boxed{\Pr(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}}$
    b. Example 1:
        i. Let's consider the same example as before, namely bound the head Pr(X>=400) where X ~ Bin(1000,0.1).
        ii. Var[x] = 1000 * 0.1 * 0.9 = 90
        iii. Pr(X >= 400) = Pr(X - 100 >= 300) <= Pr(|X - 100| >= t) <= Var[x]/t^2 = 90/300^2 = 0.001
    c. Example 2:
        i. Suppose that we toss a fair coin 100 times. What is the probability we see tails more than 60 times or less than 40 times?
        ii. Pr(X<= 40 U X>= 60) = Pr(|X-50|>= 10) <= 25/100 = ¼
3. Weak law of large numbers

a. Let X_1, X_2, … be a sequence of iid (independent identically distributed) RVs with mean μ, and standard deviation o. Consider the sum

$$S_n = X_1 + \ldots + X_n$$

b. Then as n → infinity, for all ε>0 the empirical average converges to μ in probability.

c. $$\boxed{\lim_{n\to\infty} \Pr\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) = 0.}$$ → as you increase the number of trials, the estimate from the data you obtain will be closer to the true estimate

d. Lim n → infinity Pr(|S_n/n - μ| > ε) = 0

Pr(|S_n - μ*n| > ε*n) = 0

E(S_n) = E (i=1 to i=n) (∑ X_i) = ∑ (i=1 to i=n) E(X) = n * μ

Var(S_n) = n * o^2

Pr(|S_n - μ*n| > ε*n)

Pr(|S_n - E(S_n)|| > t) <= n * o^2 / (ε^2*n^2) <= o^2/(ε^2*n)

$$\Pr\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\mathbb{V}\mathrm{ar}(S_n/n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \to 0 \text{ as } n \to \infty$$

As n → infinity, Pr(|S_n - E(S_n)|| > t) <= 0

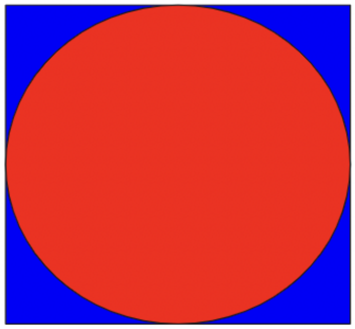
e. Example 1:
   i. Consider rolling a regular die n times. Then by WLLN we obtain that sum of the first n rolls as n grows, for any ε > 0 satisfies

$$\Pr\left(\left|\frac{S_n}{n} - \frac{7}{2}\right| > \epsilon\right) \to 0$$

4. Weak law of large numbers - Confidence intervals
   a. The proof of WLLN gives us a way to construct confidence intervals
   b. Specifically, if 0 < a < 1, we get the following confidence interval

$$\boxed{\mu - \frac{\sigma}{\sqrt{\alpha n}} \leq \frac{S_n}{n} \leq \mu + \frac{\sigma}{\sqrt{\alpha n}} \text{ wp at least } 1 - \alpha}$$

   c. P(|S_n/n - μ|>= o/sqrt(a*n))
   d. P(|S_n - μ*n|>= o*sqrt(n)/sqrt(a))
   e. P(|S_n - μ*n|>= o*sqrt(n)/sqrt(a)) <= n*o^2/(o^2*n/a) = a
5. How to estimate pi?
   a. We generate a point uniformly at random within the unit square [0,1] * [0,1]
   b. Let E be the event that the random point will fall within the circle. What is the probability of this event?

c. 

d. $$\Pr(\mathcal{E}) = \frac{\text{circle area}}{\text{area of square}} = \frac{\pi r^2}{1 \times 1} = \frac{\pi}{4}$$

e. Algorithm:
   i. We generate n random points (x_i, y_i) in the unit square [0,1] * [0,1]
   ii. We test for each point if it falls with the circle
   iii. Define S_n = #points inside the circle
   iv. $E(S_n)$ = n * $E[x_i]$ = n*pi/4
   Pi = 4 * $E(S_n)$/n
   v. Let our pi estimate be $\boxed{\tilde{\pi} = 4\frac{S_n}{n}}$
   vi. We can write the RV S_n as the sum of n independent Bernoulli variables
   vii. $$\boxed{S_n = \sum_{i=1}^{n} X_i,\ X_i = \begin{cases} 1 & \text{if i-th point inside circle} \\ 0 & \text{otherwise (o/w)} \end{cases}}$$
   viii. Var[S_n] = n*pi/4 * (1-pi/4)
   ix. Pr(|S_n/n - pi/4| >= ε) <= (pi/4)*(1-pi/4)/(nε^2) < beta (confidence interval value that we choose)
   where ε = accuracy
   x. The above equation can be written as the following by flipping it
   Pr(|S_n/n - pi/4| <= ε) >= 1 - beta (that we chose above)
   xi. Using Chebyshev's inequality, we get

$$Pr(|\frac{X_1 + \ldots + X_n}{n} - \frac{\pi}{4}| \geq \epsilon) \leq \frac{\pi/4(1-\pi/4)}{n\epsilon^2} \leq \delta$$

   xii. We can set ε = accuracy, b = confidence interval value