CS70 - Lecture 21 Notes

Name: Felix Su SID: 25794773

Spring 2016 GSI: Gerald Zhang

Distributions Review

1. **Uniform** $(U[1, \cdots, n]); m = 1, \ldots, n$
   - $\Pr[X = m] = \frac{1}{n}$
   - $E[X] = \frac{n+1}{2}$
2. **Binomial or Bernoulli** $(B(n, p)); m = 0, \ldots, n$
   - $\Pr[X = m] = \binom{n}{m} p^m (1-p)^{n-m}$
   - $E[X] = np$
3. **Geometric** $(G(p)); n = 1, 2, \ldots$
   - $\Pr[X = n] = (1-p)^{n-1} p$
   - $E[X] = \frac{1}{p}$
4. **Poisson** $(P(\lambda)); n \geq 0$
   - $\Pr[X = n] = \frac{\lambda^n}{n!} e^{-\lambda}$
   - $E[X] = \lambda$
5. **Indicator** $X = 1$ **or** 0
   - $\Pr[X = 1] = p$, $\Pr[X = 0] = (1-p)$
   - $E[X] = p$

Poisson and Queuing

(Derivation in previous notes)

1. Flip coin n times, $\Pr[H] = \frac{\lambda}{n}$
2. RV $X =$ no. of heads (Bernoulli indicator -when 1)
3. $X = B(n, \frac{\lambda}{n})$
4. Distribution of X "for large n"

- Distribution of the number of events in an interval
- The average value comes out to λ
- Cut up situation into $n \to \infty$ intervals described by Beronoulli indicators
- This means you can assume no two events occur in the same interval and there is a $\frac{\lambda}{n}$ chance the indicator is 1 in any interval

Independence Review

1. X, Y are independent if and only if:
   - $\Pr[X = x, Y = y] = \Pr[X = x]\Pr[Y = y], \forall x, y$
   - $\Pr[X \in A, Y \in B] = \Pr[X \in A]\Pr[Y \in B], \forall A, B$
2. $X, Y, Z, \ldots$ are mutually indpendent if and only if:
   - $\Pr[X = x, Y = y, Z = z, \ldots] = \Pr[X = x]\Pr[Y = y]\Pr[Z = z] \cdots, \forall x, y, z, \ldots$
   - $\Pr[X \in A, Y \in B, Z \in C, \ldots] = \Pr[X \in A]\Pr[Y \in B]\Pr[Z \in C] \cdots, \forall A, B, C, \ldots$
3. If $U, V, W, X, Y, Z, \ldots$ are all mutually independent then:
   - $f(U, V), g(W, X, Y), h(Z, \ldots), \ldots$ are mutually independent

Variance

1. Measures deviation from the mean value (Standard deviation ($\sigma(X)$) squared)
2. Use squared distance as a continuous function that you can do derivatives and other operations on.
3. To calculate intermediate value $E[X^2]$ of expectations with infinite series:
   - Calculate $E[X^2] - (1-p)E[X^2] = pE[X^2]$
   - This gives you $pE[X^2]$ in terms of $E[X]$ and a known distribution (total dist.=1)
   - Divide both sides by p
4. Variande of the constant is c that constant squared $\text{Var}[c] = c^2$

- $\text{Var}[X] = E[(X - E[X])^2]$
- $= E[X^2 - 2XE[X] + E[X]^2]$
- $= E[X^2] - 2E[X]E[X] + E[X]^2$
- $= E[X^2] - E[X]^2$

Variance of X

$$\sigma^2(X) := \text{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2 \tag{1}$$

Potential Final Question:
Give an example where this ratio $\to \infty$

$$\frac{\sigma(X)}{E[|X - E[X]|]} \tag{2}$$

Uniform Variance
Assume that $\Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2} \tag{3}$$

$$E[X^2] = \sum_{i=1}^{n} i^2 \Pr[X = i] = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{1 + 3n + 2n^2}{6} \tag{4}$$

$$\text{Var}[X] = \frac{1 + 3n + 2n^2}{6} - \frac{(n+1)^2}{4} = \frac{n^2 - 1}{12} \tag{5}$$

Geometric Distribution Variance
$X = G(p)$, $\Pr[X = n] = (1-p)^{n-1}p$ for $n \geq 1$

$$E[X] = \frac{1}{p} \tag{6}$$

$$\begin{aligned}
E[X^2] &= p + 4p(1-p) + 9p(1-p)^2 + \dots \\
(1-p)E[X^2] &= p(1-p) + 4p(1-p)^2 + 9p(1-p)^3 + \dots \\
pE[X^2] &= 2(p + 2p(1-p) + 3p(1-p)^2 + \dots) \\
&\quad - (p + p(1-p) + p(1-p)^2 + \dots) \\
&= 2E[X] - \text{Distribution} \\
pE[X^2] &= (2E[X] - 1) = (2(\frac{1}{p}) - 1) = \frac{2-p}{p}
\end{aligned}$$

$$E[X^2] = \frac{2-p}{p^2} \tag{7}$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2} \tag{8}$$

$$\sigma(X) = \frac{\sqrt{1-p}}{p} \tag{9}$$

Fixed Points Variance
Number of fixed points in a random permutation of n items
i.e. Number of students that get hw back with RV $X = X_1 + X_2 + \cdots + X_n$
X_i = indicator for ith student getting hw back

$$E[X] = 1 \tag{10}$$

$$\begin{aligned}
E[X^2] &= \sum_i E[X_i^2] + \sum_{i \neq j} E[X_i X_j] \\
&= n \times (1 \times \Pr[X_i = 1] + 0 \times \Pr[X_i = 0] = \frac{1}{n}) \\
&+ n(n-1) \times (1 \times \Pr[X_i = 1 \cap X_j = 1] + 0 \times \Pr[\text{“anything else”}] = \frac{1 \times 1 \times (n-2)!}{n!} = \frac{1}{n(n-1)})
\end{aligned}$$

$$E[X^2] = 2 \tag{11}$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = 2 - 1 = 1 \tag{12}$$

Binomial
Flip coin with $\Pr[H] = p$
$X = X_1 + X_2 + \cdots + X_n$ no. of heads (X_i, X_j are independent)
$X_i = 1$ if ith flip is heads, $X_i = 0$ otherwise (indicator RV)

$$E[X] = np \tag{13}$$

$$E[X_i^2] = 1^2 \times p + 0^2 \times (1-p) = p$$

$$\text{Var}[X_i] = p - (E[X_i])^2 = p - p^2 = p(1-p)$$

$$\text{Var}[X] = \text{Var}[X_1 + X_2 + \cdots + X_n] = n\text{Var}[X_i] = np(1-p) \tag{14}$$

Properties of Variance

> 1. $\text{Var}[cX] = c^2\text{Var}[X]$ where c is a constant
> 2. $\text{Var}[X + c] = \text{Var}[X]$ where c is a constant (shifts center)

Variance of the Sum of 2 Independent RVs

- **Thm:** If X and Y are independent, then $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$
  - Assume $E[X] = E[Y] = 0$ (You can just shift this to accountn for any other RVs)
  - $E[XY] = E[X]E[Y] = 0$ (By Independence)

$$\begin{aligned}
\text{Var}[X + Y] &= E[(X + Y)^2] \\
&= E[X^2 + 2XY + Y^2] \\
&= E[X^2] + 2E[XY] + E[Y^2] \\
&= E[X^2] + EY^2 \\
&= \text{Var}[X] + \text{Var}[Y]
\end{aligned}$$

Variance of the Sum of Multiple Independent RVs

- **Thm:** If $X, Y, Z, \ldots$ are independent, then $\text{Var}[X + Y + Z + \cdots] = \text{Var}[X] + \text{Var}[Y] + \text{Var}[Z] + \cdots$
  - Assume $E[X] = E[Y] = E[Z] = \cdots = 0$ (You can just shift this to accountn for any other RVs)
  - $E[XY] = E[XZ] = E[YZ] = \cdots = 0$ (By Independence)

$$\begin{aligned}
\text{Var}[X + Y + Z + \cdots] &= E[(X + Y + Z + \cdots)^2] \\
&= E[X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots] \\
&= E[X^2] + E[Y^2] + E[Z^2] + 0 + \cdots + 0 \\
&= \text{Var}[X] + \text{Var}[Y] + \text{Var}[Z] + \cdots
\end{aligned}$$

If X and Y are Independent

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \tag{15}$$

If $X, Y, Z, \ldots$ are Independent

$$\text{Var}[X + Y + Z + \cdots] = \text{Var}[X] + \text{Var}[Y] + \text{Var}[Z] + \cdots \tag{16}$$

Inequalities: Overview

> 1. Markov: $\Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)}$, for all a such that $f(a) > 0$ and $f : \mathbb{R} \to [0, \infty)$ is non-decreasing
>   - Bound the probability of being at least a away from the mean $E[X]$
> 2. Chabyshev: $\Pr[|X - E[X]| > a] \leq \frac{\text{Var}[X]}{a^2}$
>   - Bound probability of getting at least a away from the mean $E[X]$

Markov's Inequality Proof

- Assume $f : \mathbb{R} \to [0, \infty)$ is non-decreasing.
- $\Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)}$, for all a such that $f(a) > 0$

- Observe that: $1\{X \geq a\} \leq \frac{f(X)}{f(a)}$ because:
  - When $X > a$:
    - Left side = 1
    - Right side ¿ 1 because $f : \mathbb{R} \to [0, \infty)$ is non-decreasing
  - When $X = a$:
    - Left side = 1
    - Right side = 1
  - When $X < a$:
    - Left side = 0
    - Right side ≥ 0 because $f \geq 0$.
- Take expectation of both sides (because expectation is monotone):
  - $E[1\{X \geq a\}] \leq E[\frac{f(X)}{f(a)}] \implies \Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)}$

> **Markov's Inequality**
> $f : \mathbb{R} \to [0, \infty)$ is non-decreasing and for all a such that $f(a) > 0$
> $$\Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)} \tag{17}$$

Markov's Inequality Example $G(p)$:

- Let $X = G(p), E[X] = \frac{1}{p}, E[X^2] = \frac{2-p}{p^2}$
- Using $f(x) = x$: $\Pr[X \geq a] \leq \frac{E[X]}{a} = \frac{1}{ap}$
- Using $f(x) = x^2$: $\Pr[X \geq a] \leq \frac{E[X^2]}{a^2} = \frac{2-p}{a^2p^2}$

Markov's Inequality Example $P(\lambda)$:

- Let $X = P(\lambda), E[X] = \lambda, E[X^2] = \lambda + \lambda^2$
- Using $f(x) = x$: $\Pr[X \geq a] \leq \frac{E[X]}{a} = \frac{\lambda}{a}$
- Using $f(x) = x^2$: $\Pr[X \geq a] \leq \frac{E[X^2]}{a^2} = \frac{\lambda+\lambda^2}{a^2}$

Chebyshev's Inequality Proof

- Let $Y = |X - E[X]|, f(y) = y^2$
- Use Markov's Inequality: $\Pr[Y \geq a] \leq \frac{E[f(Y)]}{f(a)} = \frac{\mathrm{Var}[X]}{a^2}$
- Confirms that the variance measures the "deviations from the mean"

> **Chebyshev's Inequality**
> For all $a > 0$
> $$\Pr[|X - E[X]| > a] \leq \frac{\mathrm{Var}[X]}{a^2} \tag{18}$$

Chebyshev's Inequality Example $P(\lambda)$:

- Let $X = P(\lambda), E[X] = \lambda, \mathrm{Var}[X] = \lambda$
- $\Pr[|X - \lambda| > n] \leq \frac{\lambda}{n^2}$

Use Markov to get Chebyshev Bounds

- Let $X = P(\lambda), E[X] = \lambda, E[X^2] = \lambda + \lambda^2, \mathrm{Var}[X] = \lambda$

- Using Markov's with $f(x) = x^2$: $\Pr[X \geq a] \leq \frac{E[X^2]}{a^2} = \frac{\lambda+\lambda^2}{a^2}$
- If $a > \lambda$, then $X \geq a \implies X - \lambda \geq a - \lambda > 0 \implies |X - \lambda| \geq a - \lambda$
- So, for $a > \lambda$, $\Pr[X \geq a] \leq \Pr[|X - \lambda| \geq a - \lambda] \leq \frac{\lambda}{(a-\lambda)^2}$

Fraction of H's

- How likely is it that the fraction of H's differs from 50%?
- Let $X_m = 1$ if the mth flip of a fair coin is H and $X_m = 0$ otherwise
  - Any type of polling system mimics this scenario
  - $\Pr[H] = p$ Yes to poll question
  - $\Pr[T] = 1 - p$ No to poll question
- Define $Y_n = \frac{X_1+\cdots+X_n}{n}$, for $n \geq 1$
  - Ratio of people who said yes to question
- Estimate $\Pr[|Y_n - 0.5| \geq 0.1] = \Pr[Y_n \leq 0.4 \text{ or } Y_n \geq 0.6]$
  - If poll result is $\geq 10\%$ away from the mean $\implies$ mistake
- By Chebyshev: $\Pr[|Y_n - 0.5| \geq 0.1] \leq \frac{\text{Var}[Y]_n}{0.1^2} = 100\text{Var}[Y_n] = \frac{25}{n}$
  - $\text{Var}[Y_n] = \frac{1}{n^2}(\text{Var}[X_1] + \text{Var}[X_2] + \cdots + \text{Var}[X_n]) = \frac{n\text{Var}[X_i]}{n^2} = \frac{\text{Var}[X_i]}{n} \leq \frac{1}{4n}$
  - Because all X_i's are independent and the variance of a constant is its square.
  - As $n \to \infty$, $Y_n \to 0$
  - $\text{Var}[X_i] = p(1-p) \leq (.5)(.5) = \frac{1}{4}$
- For $n = 1000$, $\Pr[|Y_n - 0.5| \geq 0.1] \leq 2.5\%$

Law of Large Numbers
As $n \to \infty$, $\Pr[|X - 0.5| \geq \varepsilon] \to 0$

$$\Pr[|T_n - 0.5| \leq \varepsilon] \to 1 \tag{19}$$

Weak Law of Large Numbers

$$\Pr[|\frac{X_1 + \cdots + X_n}{n} - \mu| \geq \varepsilon] \to 0, \text{ as } n \to \infty \tag{20}$$

Weak Law of Large Numbers

Proof

- Let $Y_n = \frac{X_1+\cdots+X_n}{n}$
- $\Pr[|Y_n - \mu| \geq \varepsilon] \leq \frac{\text{Var}[Y_n]}{\mu^2} = \frac{\text{Var}[X_1+\cdots+X_n]}{n^2\varepsilon^2} = \frac{\text{Var}[X_i]}{n\varepsilon^2} \to 0, \text{ as } n \to \infty$

Summary

Variance

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2 \tag{21}$$

Fact

$$\mathrm{Var}[aX + b] = a^2 \mathrm{Var}[X] \tag{22}$$

Sum

$$X, Y, Z, \ldots \text{ mutually independent} \implies \mathrm{Var}[X + Y + Z + \cdots] = \mathrm{Var}[X] + \mathrm{Var}[Y] + \mathrm{Var}[Z] + \cdots \tag{23}$$

Markov

$f : \mathbb{R} \to [0, \infty)$ is non-decreasing and for all a such that $f(a) > 0$

$$\Pr[X \geq a] \leq \frac{E[f(X)]}{f(a)} \tag{24}$$

Chebyshev

For all $a > 0$

$$\Pr[|X - E[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2} \tag{25}$$

Weak Law of Large Numbers

$$X_m \text{ i.i.d} \implies \frac{X_1 + \cdots + X_n}{n} \approx E[X] \tag{26}$$