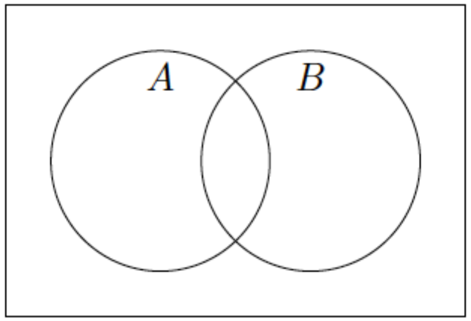
# MIDTERM REVIEW

Jian (Jessica) Gao

# Basic Probability Formulas

- What is probability
- $P(A) = (\# A) / (\# \Omega)$
- $P(A^c) = 1 - P(A)$




- $P(A \wedge B)$: "and" $= P(A)*P(B|A) = P(B)*P(A|B)$
- $P(A \vee B)$: "or" $= P(A) + P(B) - P(A \wedge B)$

- Conditional/Bayes:
- $P(A|B) = P(A \wedge B) / P(B)$
- $P(B) = P(AB) + P(A^cB)$

  $= P(B|A)*P(A) + P(B|A^c)*P(A^c)$

- Independence:
- $P(A) = P(A|B)$
- $P(AB) = P(A)*P(B)$

# Example

- How to count

- Eight cards are drawn from a well-shuffled deck of 52 cards. We say the 8 cards contain a 4 of a kind if they contain all 4 cards of a specific rank, e.g., KKKK2388 contains a 4 of a kind. What is the probability that the 8 cards contain a 4 of a kind (any kind, including the possibility of 4 of two kinds).

- Be careful of Repetition

# Distributions

| name and range | $P(k) = P(X = k)$ for $k \in$ range | mean | variance |
|---|---|---|---|
| uniform on $\{a, a+1, \ldots, b\}$ | $\frac{1}{b-a+1}$ | $\frac{a+b}{2}$ | $\frac{(b-a+1)^2 - 1}{12}$ |
| Bernoulli $(p)$ on $\{0, 1\}$ | $P(1) = p$; $P(0) = 1 - p$ | $p$ | $p(1-p)$ |
| binomial $(n, p)$ on $\{0, 1, \ldots, n\}$ | $\binom{n}{k} p^k (1-p)^{n-k}$ | $np$ | $np(1-p)$ |
| Poisson $(\mu)$ on $\{0, 1, 2, \ldots\}$ | $\frac{e^{-\mu}\mu^k}{k!}$ | $\mu$ | $\mu$ |
| hypergeometric $(n, N, G)$ on $\{0, \ldots, n\}$ | $\frac{\binom{G}{k}\binom{N-G}{n-k}}{\binom{N}{n}}$ | $\frac{nG}{N}$ | $n\left(\frac{G}{N}\right)\left(\frac{N-G}{N}\right)\left(\frac{N-n}{N-1}\right)$ |
| geometric $(p)$ on $\{1, 2, 3 \ldots\}$ | $(1-p)^{k-1}p$ | $\frac{1}{p}$ | $\frac{1-p}{p^2}$ |
| geometric $(p)$ on $\{0, 1, 2 \ldots\}$ | $(1-p)^k p$ | $\frac{1-p}{p}$ | $\frac{1-p}{p^2}$ |
| negative binomial $(r, p)$ on $\{0, 1, 2, \ldots\}$ | $\binom{k+r-1}{r-1} p^r (1-p)^k$ | $\frac{r(1-p)}{p}$ | $\frac{r(1-p)}{p^2}$ |

Differences

- \# trials fixed or not?
- Possible values
  - Start at 0 or?
  - Infinity?

Approximations to Binomials

- Poisson: n large, p small, np constant
  - Rule of thumb: sqrt(npq) < 3
- Normal: n large, p constant

Approximation to Hypergeometric

- Binomial: n << N (small chance of duplicate)

# Examples

- Find the distribution for

- The number of sixes in 15 rolls of a fair six-sided die
- The number of games of Roulette I must play until I win three times, if I only ever bet on black
- The number of Aces in 5 cards drawn from a standard deck
- 10 balls are drawn with replacement from a box with 20 blue balls and 30 non-blue balls. The number of blue balls drawn from this box
- Balls are drawn with replacement from a box with 5 red balls and 5 non-red balls until a red ball is drawn. The number of balls drawn from this box
- Customers arrive at a checkout counter randomly, at a rate of 10 per hour. The number of customers that arrive at the checkout counter in 30 minutes
- A fair 20-sided die is rolled 50 times. The face of the die shows a number greater than 15 in 30 of the rolls

# Expectation

- Method of indicators
- Tail sum:
  - for x with possible values {0, 1, …, n}
  - $E(X) = \sum_{x=1}^{\infty} P(X \geq x)$
- Familiar distribution
- Transform to other variables
- Definition $E(X) = \sum_{x} xP(X = x)$

# Examples

- Draw cards from a standard deck until three Aces have appeared. Let $X$ = number of cards drawn. Find $E(X)$

# Examples

- Suppose that four dice are rolled. Let M be the minimum of the four numbers rolled. Find E(M)

- Let X be the sum of the largest three numbers in the first four rolls. Find E(X) and Var(X)

# Examples

Suppose two teams play a series of games, each producing a winner and a loser, until one team has won two more games than the other. Let $G$ be the total number of games played. Assuming your favorite team wins each game with probability $p$, independently of the results of all previous games, find:

a. $P(G = n)$ for $n = 2, 3, \ldots$

b. $\mathbb{E}G$.

c. $Var(G)$

# Important Inequalities/Theorems

- Markov's Inequality
  - For random variable X >=0 and any a > 0

$$P(X \geq a) \leq \frac{1}{a} E(X)$$

- Chebychev's Inequality
  - For any random variable X and any k > 0

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

- Central Limit Theorem
  - For large n

Suppose the $X_i$'s are *iid* where $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$.

$$S = X_1 + X_2 + \cdots + X_n \Rightarrow S \sim N(\mu = n\mu, \sigma_S^2 = n\sigma^2)$$

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) \Rightarrow \bar{X} \sim N\left(\mu_{\bar{X}} = \mu, \sigma_S^2 = \frac{\sigma^2}{n}\right)$$

# Example

- A fair die is tossed 1200 times. Find, using an approximation:
  - The probability of getting more that 400 sixes
  - A number m such that the probability of getting between 200 - m and 200 + m sixes is approximately 95%.

# Important Math

| | | | |
|---|---|---|---|
| $\frac{1}{1-x}$ | $=\sum_{i=0}^{\infty} x^i$ | $=1+x+x^2+\cdots$ | $(-1,1)$ |
| $\log(1+x)$ | $=\sum_{i=1}^{\infty} (-1)^{i+1}\frac{1}{i}x^i$ | $=x-\frac{1}{2}x^2+\frac{1}{3}x^3-\cdots$ | $(-1,1]$ |
| $e^x$ | $=\sum_{i=0}^{\infty} \frac{1}{i!}x^i$ | $=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots$ | $(-\infty,\infty)$ |