CENG 222
Statistical Methods for Computer Engineering

Week 3

Chapter 3

Families of discrete distributions

Bernoulli distribution

- A random variable with two possible values, 0 and 1, is called a *Bernoulli variable*
- The distribution of such a r.v. is called the *Bernoulli distribution*
- Any random experiment with a binary outcome is called a *Bernoulli trial*
- Generic outcome names: *successes* and *failures*

Not equally likely outcomes

- In general, $f(1) = f(0) = 0.5$ does NOT hold when the binary outcomes are not equally likely
- If $f(1) = p$, what is $E(X)$ and $Var(X)$?

What about “non 0-1”, binary outcomes?

- Example:
  - What if the two possible outcomes are 5 and 9 with $f(5) = 0.3$ and $f(9) = 0.7$?
  - What is the expected value?

What about “non 0-1”, binary outcomes?

- Example:
  - What if the two possible outcomes are 5 and 9 with $f(5) = 0.3$ and $f(9) = 0.7$?
  - What is the expected value?
  - It is just a shifted and rescaled standard Bernoulli trial.
    - $X = 4B + 5$
    - $E(X) = E(4B + 5) = 4E(B)+5 = 4 \cdot 0.7+5 = 7.8$

Binomial distribution

- Number of successes in a sequence of independent Bernoulli trials
  - n: number of trials
  - p: probability of success
- $f_x(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}$
- Expected value and variance:
  - A binomial variable X is a sum of n independent Bernoulli trials.
  - $E(X) = np$, $Var(X) = npq$

Using distribution tables

- Table A2, *cdf* of Binomial distribution
- *pdf* can be obtained by difference of two consecutive entries
- Example 3.16
- Example 3.17
  - Using *binocdf*(x,n,p) function of MATLAB

Geometric distribution

- The number of Bernoulli trials needed to get the first success
- The support is the set of integers [1..∞]
- $f_x(x) = P(X = x) = pq^{x-1}$
- The support is unbounded
  - Check that $\sum_x f_x(x) = \sum_{x=1}^{\infty} p(1-p)^{x-1} = 1$
- Expected value and variance:
  - $E(X) = 1/p$, $Var(X) = (1-p)/p^2$

Geometric distribution

- Example 3.20 St. Petersburg Paradox
- Gambling with a guaranteed strategy to win a desired amount
  - Even when p is less then 0.5!
  - Start with the desired amount
  - Double betting amount every time you loose
  - Stop when you win the first time
  - E.g if p=0.2 the expected number of bets to win is 5!

Geometric distribution

- So what's the paradox?
- What is the amount of money, Y, needed to follow the strategy?
  - $Y = D2^{X-1}$ where D is the desired amount and X is the number of bets needed to win.
  - $E(Y)$ = infinity when $p \leq 0.5$ (the paradox)

Negative Binomial distribution

- In a sequence of independent Bernoulli trials, the number of trials needed to obtain k successes
  - It can be considered as the *inverse* of the Binomial, where, we now fix the number of successes and count the number of trials n to reach that number of successes
- It is a generalization of the Geometric distribution

Negative Binomial distribution

- $f_x(x) = P(X = x) = \binom{x-1}{k-1} p^k q^{x-k}$
- Expected value and variance:
  - A negative binomial variable X is a sum of k independent Geometric variables.
  - $E(X) = k/p$, $Var(X) = k(1\text{-}p)/p^2$
- Example 3.21
  - $k = 12$, $p = 0.95$, $P(X > 15) = ?$
  - $P(X>15) = 1 - F_X(15)$
  - Can be solved by using the Binomial distribution with $n = 15$, $p = 0.95$, $P(Y < 12) = F_Y(11)$.

Poisson distribution

- The number of rare events occurring within a fixed period of time
- It has a single parameter
  - λ : frequency, average number of events
  - $f_x(x) = e^{-\lambda} \frac{\lambda^x}{x!}$
  - $E(X) = \lambda$, $Var(X) = \lambda$
- Example 3.22

Poisson approximation of Binomial distribution

- Poisson distribution can be used to approximate Binomial distribution when n is large and p is small
  - E.g., $n \geq 30$ and $p \leq 0.05$
  - $np = \lambda$
- Example 3.25 The Birthday Problem