Guidelines for Solving Probability Problems

CS 1538: Introduction to Simulation

1 Steps for Problem Solving

Suggested steps for approaching a problem:

1. Identify the distribution – What distribution does the situation seem to follow?
2. Identify the item to solve for – Usually this is a probability, but it can be an expected value or something else. If it's a probability, is it for a specific outcome or a range of outcomes?
3. From the problem, extract values for the parameters of the distribution.
4. Write out the appropriate equation, plug in the values, and solve.

2 Identifying Distributions

The guidelines below can be helpful in determining the correct distribution to use.

- Do you have a fixed number of binary trials? – Binomial
  - If that fixed number is 1, then use Bernoulli (which is really just Binomial with n=1)
- Are you waiting for the *first* success in a series of binary trials? – Geometric

- Are you waiting for the k^{th} success in a series of binary trials? – Negative Binomial
- Are you dealing with the arrival or occurrence of events?
  - If you're being asked about the number of arrivals/occurrences in a time period – Poisson
  - If you're being asked about the time between one event and the next – Exponential
  - If you're being asked about the time until k events – Erlang

3 General Equations

3.1 Expected Value

The expected value, or mean, is represented by μ. It is calculated by:

- Discrete: $E[X] = \sum_i p(x_i) * x_i$
- Continuous: $E[X] = \int_{-\infty}^{\infty} f(x) * x dx$

3.2 Variance

The variance represents how different each random variable is from the mean. It's the expectation of how much difference there is between each possible outcome and the mean.

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

3.3 Probability

- $\Pr(a < X < b) = \Pr(X > a, X < b) = \Pr(X < b) - \Pr(X < a)$
- $\Pr(A|B) = \Pr(A, B) / \Pr(B)$

4 Distributions

Below are the distributions we've talked about in class.

Note: In the equations below, q is defined to be the probability of failure and will have the value $q = 1 - p$.

4.1 Bernoulli

Only one trial and it has two possible outcomes.
Examples:

- A coin toss

Parameters:

- p: Probability of success; range: [0, 1]

Formulas:

- $\Pr(X = x) = p$
- $E[X] = p$
- $Var(X) = p(1 - p)$

4.2 Binomial

Bernoulli trials are repeated a set number (n) times.
Examples:

- Flipping a coin n times with a bias of p% of being heads

Parameters:

- n: Number of trials; positive integer
- p: Probability of success in one trial; range: [0, 1]

Formulas:

- $\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$

- CDF: $\Pr(X \leq x) = \sum_{i=0}^{\lfloor x \rfloor} \Pr(X = i)$
- $E[X] = np$
- $Var(X) = np(1-p)$

4.3 Geometric

Bernoulli trials are repeated until a trial is successful, then stop.
Examples:

- Flipping a coin until you get heads.
- What's the probability that a product will fail after k uses?

Parameters:

- p: Probability of success in one trial; range: [0, 1]

Formulas:

- $\Pr(X = x) = \begin{cases} q^{x-1}p & x = 1, 2, ... \\ 0 & \text{otherwise} \end{cases}$
- CDF: $\Pr(X \leq x) = \sum_{i=0}^{\lfloor x \rfloor} \Pr(X = i)$
- $\Pr(X > x) = q^x$
- $E[X] = 1/p$
- $Var(X) = q/p^2$

4.4 Negative Binomial

Bernoulli trials are repeated until k trials are successful, then stop.
Examples:

- Flipping a coin until you get 4 heads.

Parameters:

- p: Probability of success in one trial; range: [0, 1]

- k: Number of successes needed before stopping; positive integer

Formulas:

- $\Pr(X = x) = \begin{cases} \binom{x-1}{k-1} q^{x-k} p^k & x = k, k+1, k+2, ... \\ 0 & \text{otherwise} \end{cases}$
- CDF: $\Pr(X \leq x) = \sum_{i=0}^{\lfloor x \rfloor} \Pr(X = i)$
- $E[X] = k/p$
- $Var(X) = kq/p^2$

4.5 Poisson

The probability of x events occurring in a fixed interval of time if these events occur with a known average rate (α) and independently of the time since the last event.
Examples:

- A person flips a coin once every 5 seconds. They observe that they get heads about 6 times every minute (α). What's the probability of x heads in the next hour?
- Port Authority reports that, for the middle of the day during weekdays, the 71A runs every 15 minutes (α). What is the probability of 4 buses in an hour?
- Photons arrive at a telescope at a rate of 1 every 300 μs (α). What is the probability that 4000 would arrive in one minute?
- A certain (possible under-used) traffic light has 1 car show up every 20 seconds (α). What's the probability that 10 cars will show up in 20 seconds?

Parameters:

- α: Average rate of occurrence for the event of interest; $\alpha > 0$

Formulas:

- $\Pr(X = x) = \begin{cases} \frac{e^{-\alpha} \alpha^x}{x!} & x = 0, 1, ... \\ 0 & \text{otherwise} \end{cases}$

- CDF: $\Pr(X \leq x) = \sum_{i=0}^{\lfloor x \rfloor} \Pr(X = i)$
- $E[X] = \alpha$
- $Var(X) = \alpha$

Let $\alpha = \lambda t$, with λ being the arrival rate within a time interval $[0, t]$. With this transformation, we are able to get time explicitly into our formulas.

4.6 Uniform

Every value in the state space is equally likely.
Examples:

- It is known that the time to complete an oil change at a certain dealership is random and takes between 35 and 45 minutes. If you take your car to this dealership for an oil change, what is the probability that it takes between 40 and 42 minutes?

Parameters:

- a: lower bound of range of possible values
- b: upper bound of range of possible values

Formulas:

- $f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$
- CDF: $F(x) = \frac{x-a}{b-a}$
- $E[X] = \frac{b+a}{2}$
- $Var(X) = \frac{(b+a)^2}{12}$

4.7 Normal

Commonly found in natural and behavioral sciences for measuing variations around a population average. This is the "bell-shaped curve" distribution. Examples:

- On average, an ocean-going vessel requires 12 hours to load, with a variance of 4 hours^2. Assume that loading times are normally distributed. What is the probability of requiring less than 10 hours to load?

- Suppose that wait time (in minutes) follows a Normal distribution with mean 15 and variance 9. What is the probability of having to wait between 14 and 17 minutes?

Parameters:

- μ: mean
- σ^2: variance

Formulas:

- $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- CDF: $F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(t)dt$
- $E[X] = \mu$
- $Var(X) = \sigma^2$
- Transformation from Normal Distribution to Standard Normal Distribution: $z = \frac{x-\mu}{\sigma}$.

Tips:

- Problems requiring the Uniform distribution can appear similar to those requiring the Normal distribution. How can you tell the difference? If the problem gives a range and values cannot be outside of that range, then it's the Uniform distribution. If it gives an average and some measure of variability (i.e. variance or standard deviation), then use the Normal distribution.

4.8 Exponential

Used to describe the time before a single Poisson event.
Examples:

- Time until a radioactive particle decays.
- Port Authority reports that, for the middle of the day during weekdays, the 71A runs every 15 minutes (λ). What is the probability of a wait time more than 15 minutes?

Parameters:

- λ: arrival rate

Formulas:

- $f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$
- CDF: $F(x) = \Pr(X \leq x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$
- $E[X] = 1/\lambda$
- $Var(X) = 1/\lambda^2$
- Memoryless: $\Pr(X > s + t | X > s) = \Pr(X > t)$

Tips:

- The Poisson distribution is concerned with the number of events in a given time. The Exponential distribution is concerned with the time between two events.

4.9 Erlang

Used to describe the time before k Poisson events.
Examples:

- Time until k radioactive particle decays.

- Port Authority reports that, for the middle of the day during weekdays, the 71A runs every 15 minutes (λ). If I'm watching buses go by, what is the probability that I'll have to wait 20 minutes to see see 2 buses go by?

Parameters:

- k: number of events
- θ: rate of all events

Formulas:

- $f(x) = \begin{cases} \frac{\beta\theta}{\Gamma(\beta)}(\beta\theta x)^{\beta-1}e^{-\beta\theta x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$
- CDF: $F(x) = \Pr(X \leq x) = \begin{cases} 1 - \sum_{i=0}^{k-1} \frac{e^{-k\theta x}(k\theta x)^i}{i!} & x > 0 \\ 0 & \text{otherwise} \end{cases}$
- $E[X] = \frac{1}{\theta}$
- $Var(X) = \frac{1}{\beta\theta^2}$
- Arrival rate of one item: $\lambda = k\theta$