Review of Probability Distributions

CS1538: Introduction to Simulations

Some Well-Known Probability Distributions

- Bernoulli
- Binomial
- Geometric
- Negative Binomial
- Poisson
- Uniform
- Exponential
- Gamma
- Erlang
- Gaussian/Normal

Relevance to simulations:

- Need to use distributions that are appropriate for our problem
- The closer the chosen distribution matches the distribution in reality, the more accurate our model
- Why not always make a user-defined distribution specific to our problem?

Discrete Distributions

- Bernoulli
- Binomial
- Geometric
- Negative Binomial
- Poisson

The Bernoulli Distribution

- Bernoulli Trial: an experiment with only two possible outcomes: Success or Failure
- The probability of success is p (where $0 \leq p \leq 1$)
  - Let q be the probability of failure. *What is q in terms of p?*
- Examples?

The Bernoulli Distribution

- Let B be a random variable over the outcome of the experiment:
  - B = 1 for a success; B = 0 for a failure
  - Expectation of B
    - $E[B] = 1*p + 0*(1-p) = p$
  - Variance of B
    - $Var(B) = E[B^2] - (E[B])^2 = (1^2 p + 0^2 (1-p)) - p^2 = p - p^2 = p(1-p)$

Binomial Distribution

- Experiment: Repeat Bernoulli Trial for n times
  - Good for determining the probability of getting k defective items in a batch size of n
  - Let random variable X be the number of successes
    - Note: the order doesn't matter
    - For example: suppose we toss a biased coin 3 times (with heads=success).
      - X(HHT)=2; X(THH)=2; X(HTH)=2

Binomial Distribution

- The probability mass function for X:

$$p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0,1,\ldots,n \\ 0, & otherwise \end{cases}$$

- $E[X] = E[B_1 + B_2 + \ldots B_n] = \Sigma\, E[B_i] = np$
- $Var(X) = Var(B_1 + B_2 + \ldots B_n) = \Sigma\, Var(B_i) = np(1\text{-}p)$

- Note: $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

Geometric Distribution

- Keep on repeating Bernoulli trials until successful
- Let r.v. X be the number of trials until the first success
- The probability mass function for X:

$$p(x) = \begin{cases} q^{x-1}p, & x = 1,2,\cdots \\ 0, & otherwise \end{cases}$$

  - What is q?

Geometric Distribution

- $E[X] = \Sigma\, xq^{(x-1)}p = p\, \Sigma\, xq^{(x-1)} = 1/p$
  - We know that $\int \Sigma\, xq^{(x-1)}\, dq = \Sigma\, q^x$ where $x = 1,2,\ldots$
  - We also know that $\Sigma\, q^y$ for $y = 0,1,2\ldots = 1/(1-q)$
  - So $\int \Sigma\, xq^{(x-1)}\, dq = 1/(1-q) - 1 = q/1-q$
  - $d/dq\; q/(1-q) = q/(1-q)^2 + 1/(1-q) = 1/(1-q)^2 = 1/p^2$
  - So $p\, \Sigma\, xq^{(x-1)} = d/dq \int \Sigma\, xq^{(x-1)}\, dq = p\,(1/p^2) = 1/p$

- $Var(X) = q/p^2$

Example Question

- Suppose a product has a p% chance of failure. What's the chance that it will still be working after 3 uses?

Example Question

- What's the chance that the product is still working after 7 uses, given that it works after 3 uses?

Geometric Distribution Properties

- $Pr(X > t) = q^t$

- Memoryless
  - $Pr(X > s+t \mid X > s) = Pr(X > t)$

Negative Binomial Distribution

- Keep on running Bernoulli Trials until we get k successes.
  - Like Geometric, but now k successes instead of 1 success
- Let r.v. X be the total number of trials
- The probability mass function for X:

$$p(x) = \begin{cases} \binom{x-1}{k-1} q^{x-k} p^{k}, & x = k, k+1, k+2 \cdots \\ 0, & otherwise \end{cases}$$

- Expectation: E[X] = k/p
- Var(X) = kq/p^2

Poisson Distribution

- Computes the probability of the number of events that may occur in a period, given the rate of occurrence in that period
- Good for modeling arrivals
- Probability mass function:

$$p(x) = \begin{cases} \dfrac{e^{-\alpha}\alpha^{x}}{x!}, & x = 0,1,\cdots \\ 0, & otherwise \end{cases}$$

  - Where α is a fixed value that must be positive. It represents the average rate of the event of interest occurring.

Example

- The Prussian Cavalry/Horse study [Bortkiewicz, 1898; cf. Larsen&Marx]
  - 10 cavalry corps monitored over 20 years.
  - X = number of fatality due to horse kicks

x = # of deaths	Observed # of corps-years in which x fatality occurred	Expected # of corps-year using Poisson
0	109	
1	65	
2	22	
3	3	
4	1	
Total	200	

Poisson Distribution

- Cumulative distribution function:

$$F(x) = \sum_{i=0}^{x} \frac{e^{-\alpha}\alpha^{i}}{i!}$$

  - Note that when x → ∞, F(x) → 1

$$\sum_{i=0}^{\infty} \frac{e^{-\alpha}\alpha^{i}}{i!} = e^{-\alpha} \sum_{i=0}^{\infty} \frac{\alpha^{i}}{i!} = e^{-\alpha} e^{\alpha} = 1$$

- $E[X] = \alpha$
- Var(X) = α

Example

- Suppose that I see students at the rate of five per office hours. What is the chance that I'll see 4 students at office hours today?

- What is the chance that I'll see less than 3 students at office hours today?

- What is the chance that I'll see at least 3 students at office hours today?

Poisson: Relationship to Binomial

- **Theorem**: Let α be a fixed number and n be an arbitrary positive integer. For each nonnegative integer x,

$$\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{e^{-\alpha}\alpha^x}{x!}$$

where $p = \alpha/n$

- This makes it easier to calculate the binomial distribution, especially for large n's

Poisson Arrival Process

- Recall α represents average rate of occurrence
  - How do we represent time more explicitly?

- Let N(t) be a random variable that represents the number (positive integer) of events that occurred in time [0, t]
  - We want to know Pr(N(t) = n)
  - If we guarantee a few properties, we can use the Poisson Distribution

Poisson Arrival Process

- Assumptions:
  - Arrivals occur one at a time
  - Arrival rate (λ) does not change over time
    - We'll see how to relax this
  - The number of arrivals in given period are independent of each other
- We can now rewrite the Poisson Distribution:
  - $\alpha = \lambda t$

$$P[N(t)=n]=\begin{cases}\dfrac{e^{-\lambda t}(\lambda t)^n}{n!}, & n=0,1,\cdots \\ 0, & otherwise\end{cases}$$

- How does this affect the expectation and variance?

Properties of the Poisson Arrival Process

- Random Splitting
  - Consider a Poisson Process $N(t)$ with rate λ. Assume that arrivals can be divided into two groups, A and B with probability p and (1-p), respectively
  - N_A is a Poisson Process with rate λp and N_B is a Poisson Process with rate $\lambda(1-p)$
  - $N(t) = N_A(t) + N_B(t)$

Properties of the Poisson Arrival Process

- Pooled Process
  - Consider two Poisson Processes $N_1(t)$ and $N_2(t)$, with rates λ_1 and λ_2
  - The sum of the two processes is also a Poisson Process; it has a rate of $\lambda_1 + \lambda_2$
  - Pooling can be used in situations where multiple arrival processes feed a single queue

Discrete Probability Summary

- Binomial Distribution
- Geometric Distribution
- Negative Binomial Distribution

These describe experiments based on repeated Bernoulli Trials

- Poisson Distribution

Doesn't have an easy to describe underlying structure, but seems to be a good fit for many real data sets

Continuous Random Variables

- Random variable X is continuous if its *range space* is an interval or a collection of intervals
- There exists a non-negative function f(x), called the probability density function, such that for any set of real numbers,
  - f(x) >= 0 for all x in the range space
  - $\int_{rangespace} f(x)dx = 1$ (i.e., the total area under f(x) is 1)
  - f(x) = 0 for all x not in the range space

Note that f(x) does **not** give the probability of X = x

- Unlike the pmf for discrete random variables

Continuous Random Variables

- The probability that X lies in a given interval [a,b] is

$$P(a \le X \le b) = \int_a^b f(x)dx$$

  - aka "area under the curve"
  - Note that for continuous random variables,
    Pr(X = x) = 0 for any x
  - Consider the probability of x within a (very small) range
- The cumulative distribution function (cdf), F(x) is now the integral from -∞ to x or

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

    - This gives us the probability up to x

Continuous Random Variables

- Expected Value for a continuous random variable
  - Similar to the discrete case, except that we integrate instead of summing

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

- Variance: same formulation as its discrete counterpart (though calculating $E[X^2]$ will involve integrals again).

$$Var(X) = E[X^2] - (E[X])^2$$

Uniform Distribution over range [a,b]

- Probability density function:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & if\ a \le x \le b \\ 0 & otherwise \end{cases}$$

- Cumulative distribution function:

$$F(x) = \int_a^x f(y)dy = \int_a^x \frac{1}{b-a}dy = \frac{x-a}{b-a} \quad if\ a \le x \le b$$

  - What about F(x) when x < a or x > b?

Uniform Distribution over range [a,b]

- **Expectation**

$$E(X) = \int_a^b \frac{x}{b-a}\,dx = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)} = \frac{b+a}{2}$$

  - What does the expected value of a discrete uniform distribution look like?

- **Variance**

$$Var(X) = E[X^2] - (E[X])^2 = \int_a^b \frac{x^2}{b-a}\,dx - (E[X])^2$$

$$= \left.\frac{x^3}{3(b-a)}\right|_a^b - \frac{(b+a)^2}{4} = \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4}$$

$$= \frac{4(b^3-a^3) - 3(b+a)^2(b-a)}{12(b-a)} = \frac{(b+a)^2}{12}$$

Normal Distribution

- Probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty$$

where we supply the mean and variance:

  - μ: mean
  - σ: square-root of variance
  - The normal distribution is also denoted as: $N(\mu,\sigma^2)$
- Some of its properties:

$$\lim_{x\to-\infty} f(x) = 0 = \lim_{x\to\infty} f(x)$$

$$f(\mu+x) = f(\mu-x)$$

$$\max(f(x)) = f(\mu)$$

Normal Distribution

- F(x) = P(X≤x) = $\int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt$
  - Doesn't have a closed form.
  - Can use numerical methods, but want to avoid evaluating integrals for each pair (μ, σ^2)
    - Transform to standard normal distribution
  - Let $z = (t-\mu)/\sigma$ then we can rewrite the above as:
    - $F(x) = P(X \le x) = P(Z \le (x-\mu)/\sigma)$

$$= \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z)^2\right] dz$$

CDF table for N(0,1)

- There are a few variants. On the website:

Z	.00	.01	.02	.03
-3.9	.00005	.00005	.00004	.0000
-3.8	.00007	.00007	.00007	.0000
-3.7	.00011	.00010	.00010	.0001
-3.6	.00016	.00015	.00015	.0001
-3.5	.00023	.00022	.00022	.0002
-3.4	.00034	.00032	.00031	.0003
-3.3	.00048	.00047	.00045	.0004
-3.2	.00069	.00066	.00064	.0006

- A typical table from a stats book lets you specify z to 2 significant digits: ($z = column_1 + row_1$)

Example (from textbook 5.21)

- Suppose we have a normal distribution such that:

  X is a r.v. from N(50, 9)

  - What is the chance that $X \leq 56$?
    - We wish to compute F(56) (i.e., $P(X \leq 56)$).

Exponential Distribution

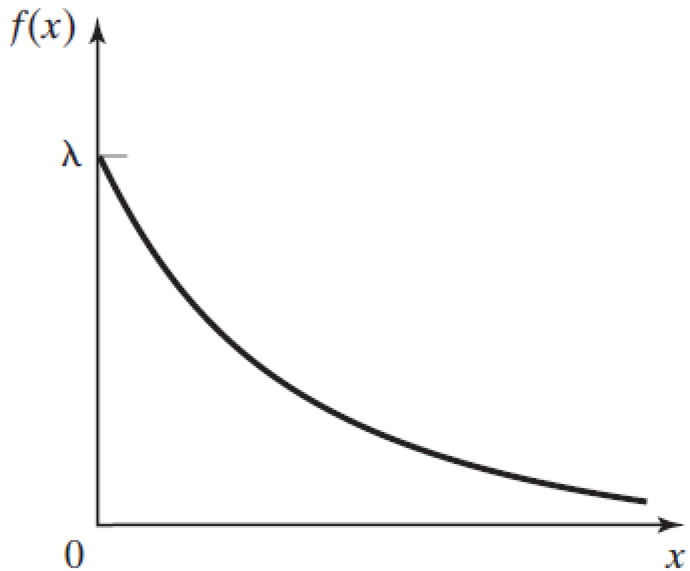
- Models
  - Interarrival times
  - Service times
  - Lifetime of a component that fails instantaneously
- Parameter λ indicates rate

Exponential Distribution

- Probability density function:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

  - where $\lambda > 0$ is a parameter that we supply
  - Since the exponent is negative, the pdf will decrease as x increases

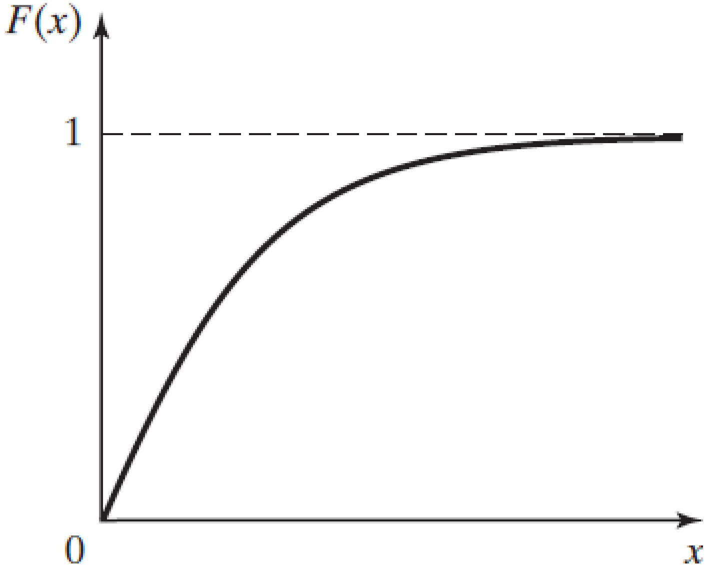


Exponential Distribution

$$F(x)=\begin{cases}0, & x<0\\ \int_0^x \lambda e^{-\lambda t}dt = 1-e^{-\lambda x}, & x\geq 0\end{cases}$$

$$E[X]=\frac{1}{\lambda}$$

$$Var(X)=\frac{1}{\lambda^2}$$



Exponential Distribution: λ

- λ represents a rate: number of occurrences per time unit
- x: amt. of time it took for some occurrence to take place
- F(x): prob. that the event happened during interval [0,x]
- 1-F(x): prob. that the event doesn't happen until after x

Example 5.17

Suppose lifespan of a type of lamp is exponentially distributed with failure rate $\lambda = 1/3000$ hrs. What's the chance a particular lamp will beat the average?

- Has a relationship to the Poisson Arrival Process

Exponential Distribution

- Like the geometric distribution, the exponential distribution is memoryless
  - $P(X > s+t \mid X > s) = P(X > t)$
  - The proof is similar to how we showed that the geometric distribution is memoryless

Relationship between Poisson Distribution and Exponential Distribution (1/3)

- Events occur at a rate of λ
- Poisson: chance of n events taking place is given by:

  $$Pr[N(t)=n] = e^{-\lambda t} (\lambda t)^n/n!$$

- Let E be the first occurrence of an event since the start of the clock at 0; let t represent the time when E happens
- Want to find: distribution for t
- Look for CDF F(t) – the probability that E happened some time within the interval [0,t]
  - Then we can take the derivative of F(t) to find f(t)
  - Want to show: F(t) fits Exponential

Relationship between Poisson Distribution and Exponential Distribution (2/3)

- F(t) = The chance that E occurs in some interval [0,t]
  = 1 – Pr(E occurred after t) = 1 – Pr(nothing took place in [0,t])
- *Nothing took place in [0,t]* relates counting discrete event occurrences with measuring continuous time duration
  - Pr(nothing took place in [0,t]) = Pr[N(t)=0]
  - F(t) = 1 – Pr(E occurred after t) = 1 – Pr[N(t) = 0] = $1 - e^{-\lambda t}$
    which is the CDF for exponential distribution
- So the duration between the 0^{th} and 1^{st} event of a Poisson arrival process follows an exponential distribution

Relationship between Poisson Distribution and Exponential Distribution (3/3)

- More generally, for any two events E_i and E_j that follows the Poisson arrival process, the duration between them also follows the exponential distribution

- Let i be the time when E_i occurs and j be the time when E_j occurs
- We can use the same analysis as before (just align i with 0 and j with t in the previous case)
  - This is because Poisson arrival process assumes constant rate
  - Also makes sense – exponential distribution is memoryless

Exercises with Exponential

- Rate: 0.5 arrivals per minute
- $F(t) = 1 - e^{-0.5t}$ = probability of time elapse between arrivals
  - What's the probability that the next customer shows up before 30 seconds have passed?
  - … within 1 minute?
  - … within 2 minutes?
  - What's the chance that a customer shows up between 5 and 7 minutes?

Relationship between Gamma Distribution and Poisson Distribution

- Suppose some event occurs over time interval of length x at the average rate of λ per unit time. How long would it take for r events to happen?
  - Divide up x into small independent subintervals such that the chance of more than one event happening in the subinterval is negligibly small
  - Let W be a r.v. counting the # of occurrences of the event in the total duration [0,x]. W is a Poisson r.v. w/ params λx
  - $F_X(x) = Pr(X \le x) = 1 - Pr(X > x)$

    $= 1 - F_W(r-1) = 1 - \sum_{k=0}^{r-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!}$
  - $f_X(x) = F'_X(x) = \frac{\lambda^r}{(r-1)!} x^{r-1} e^{-\lambda x}$

The Gamma function:

- If r can be non-integer, then we need to replace (r-1)! with a continuous function of r.
- We'll call this function *Gamma* of r: $\Gamma(r)$
  - For any real number r > 0, the gamma function $\Gamma(r)$ is:

$$\Gamma(\beta) = \int_0^\infty x^{\beta-1} e^{-x} dx$$

    - $\Gamma(1) = 1$
    - $\Gamma(1/2) = \text{sqrt}(\pi)$
    - $\Gamma(r+1) = r\,\Gamma(r)$ for any positive real r
    - $\Gamma(r+1) = r!$ if r is a nonnegative integer
    - $\binom{n+r-1}{n} = \frac{\Gamma(n+r)}{\Gamma(n+1)\Gamma(r)}$
    - $\frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} = \int_0^1 u^{r-1}(1-u)^{s-1} du$

Gamma Distribution (general form)

- Useful for when waiting times between events is relevant (e.g. waiting time between Poisson events)
- Let X be a random variable such that

$$f_X(x) = \begin{cases} \frac{\beta\theta}{\Gamma(\beta)}(\beta\theta x)^{\beta-1}e^{-\beta\theta x}, & x > 0, \beta > 0 \\ 0, & otherwise \end{cases}$$

  - β is referred to as a shape parameter
  - θ is referred to as a scale parameter

- E[X] = $\int_0^\infty x\, f(x)\, dx \; = \frac{1}{\theta}$
- Var(X) = $\frac{1}{\beta\theta^2}$
- CDF: $F_X(x) = \int_0^x f(t)\, dt = 1 - \int_x^\infty f(t)dt$

Erlang Distribution

- When β is an arbitrary positive integer, the Gamma Distribution is called the Erlang Distribution of order k (k= β)
- Can simplify the CDF to:
  - $$F(X) = \begin{cases} 1 - \sum_{i=0}^{k-1} \frac{e^{-k\theta x}(k\theta x)^i}{i!} & x > 0 \\ 0 & x \leq 0 \end{cases}$$
  - This is the sum of Poisson terms with mean $\alpha = k\theta x$
  - k is the number of events
  - θ is the rate of the collection of events
  - $\lambda = k\theta$ is the rate of one event (events per unit time)

Erlang Example

- Suppose a node in a network does not transmit until it has accumulated 5 messages in its buffer. Suppose messages arrive independently and are exponentially distributed with a mean of 100 ms between messages.
  - Suppose a transmission was just made; what's the probability that more than 552 ms will pass before the next transmission?