Continuous Random Variables

- Recall: For *discrete* random variables, only a *finite* or *countably infinite* number of possible values with positive probability.
  - ➢Often, there is interest in random variables that can take (at least theoretically) on an *uncountable* number of possible values, e.g.,

    the weight of a randomly selected person in a population,

    the length of time that a randomly selected light bulb works,

    the error in experimentally measuring the speed of light.
  - ➢Example (Uniform Spinner, LNp.2-14):
    - $\Omega = (-\pi, \pi]$
    - For $(a, b]\subset\Omega$, $P((a, b]) = b-a/(2\pi)$
    - Consider the random variables:

      $X: \Omega \to \mathbb{R}$, and $X(\omega) = \omega$ for $\omega\in\Omega$,

      $Y: \Omega \to \mathbb{R}$, and $Y(\omega) = tan(\omega)$ for $\omega\in\Omega$.

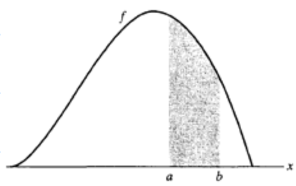
      Then, X and Y are random variables that takes on an uncountable number of possible values.

- Notice that:
  - $P_X(\{X = x\})=0$, for any $x \in \mathbb{R}$,
  - But, for $-\pi \leq a<b\leq\pi$,

    $$P_X(\{X \in (a, b]\})=P((a, b]) = b-a/(2\pi) > 0.$$

  Q: Can we still define a probability mass function for X? If not, what can play a *similar* role like pmf for X?

- Probability Density Function and Continuous Random Variable
  - ➢Definition. A function f: $\mathbb{R}\to\mathbb{R}$ is called a probability density function (pdf) if
    1. $f(x) \geq 0$, for all $x\in(-\infty, \infty)$, and
    2. $\int_{-\infty}^{\infty} f(x)dx = 1.$



  - ➢Definition: A random variable X is called *continuous* if there exists a pdf f such that for any set B of real numbers

    $$P_X(\{X \in B\}) = \int_B f(x)\, dx.$$

    - For example, $P_X(a \leq X \leq b) = \int_a^b f(x)dx.$

➢Theorem. If f is a pdf, then there must exist a continuous random variable with pdf f.

➢Some properties

- $P_X(\{X=x\}) = \int_x^x f(y)dy = 0$ for any $x \in \mathbb{R}$
- It does not matter if the intervals are open or close, i.e.,

$$P(X \in [a,b]) = P(X \in (a,b]) = P(X \in [a,b)) = P(X \in (a,b)).$$

- It is important to remember that the value a pdf $f(x)$ is NOT a probability itself
- It is quite possible for a pdf to have value greater than 1
- **Q**: How to interpret the value of a pdf $f(x)$? For small dx,

$$P\left(x - \frac{dx}{2} \le X \le x + \frac{dx}{2}\right) = \int_{x-\frac{dx}{2}}^{x+\frac{dx}{2}} f(y)dy \approx f(x)\cdot dx.$$

$\Rightarrow$ $f(x)$ is a measure of how likely it is that X will be near x

➢We can characterize the distribution of a continuous random variable in terms of its

1. Probability Density Function (pdf)
2. Cumulative Distribution Function (cdf)
3. Moment Generating Function (mgf, Chapter 7)

- Relation between the pdf and the cdf

➢Theorem. If F_X and f_X are the cdf and the pdf of a continuous random variable X, respectively, then

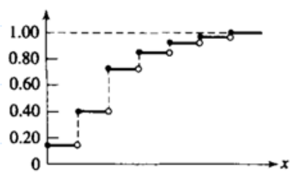
- $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(y)dy$ for all $-\infty < x < \infty$
- $f_X(x) = F'_X(x) = \frac{d}{dx}F_X(x)$ at continuity points of f_X

➢Some Notes

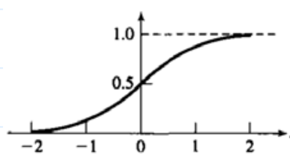
- For $-\infty \le a < b \le \infty$

$$P(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx.$$

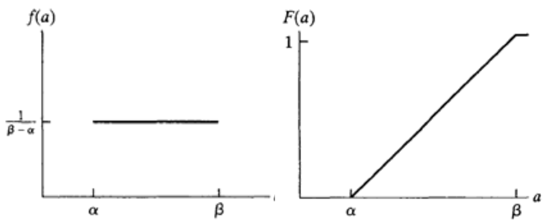
- The cdf for continuous random variables has the same interpretation and properties as in the discrete case



- The only difference is in plotting F_X. In the discrete case, there are *jumps*. In the continuous case, F_X is a *continuous* non-decreasing function.



➢Example (Uniform Distributions)



- If $-\infty<\alpha<\beta<\infty$, then

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \text{if } \alpha < x \le \beta, \\ 0, & \text{otherwise,} \end{cases}$$

  is a pdf since

  1. $f(x) \ge 0$ for all $x \in \mathbb{R}$
  2. $\int_{-\infty}^{\infty} f(x)dx = \int_{\alpha}^{\beta} \frac{1}{\beta-\alpha} dx = \frac{1}{\beta-\alpha}(\beta-\alpha) = 1.$

- Its corresponding cdf is

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 0, & \text{if } x \le \alpha, \\ \frac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha < x \le \beta, \\ 1, & \text{if } x > \beta. \end{cases}$$

- (exercise) Conversely, it can be easily checked that F is a cdf and $f(x)=F'(x)$ except at $x=\alpha$ and $x=\beta$ (Derivative does not exist when $x=\alpha$ and $x=\beta$, but it does not matter.)

- An example of Uniform distribution is the r.v. X in the Uniform Spinner example where $\alpha=-\pi$ and $\beta=\pi$.

• Transformation

➢ **Q**: $Y=g(X)$, how to find the distribution of Y?

- Suppose that X is a continuous random variable with cdf F_X and pdf f_X.
- Consider $Y=g(X)$, where g is a strictly monotone (increasing or decreasing) function. Let R_Y be the range of g.
- Note. Any strictly monotone function has an inverse function, i.e., g^{-1} exists on R_Y.

➢The cdf of Y, denoted by F_Y

1.Suppose that g is a strictly increasing function. For $y \in R_Y$,

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(g(X) \le y) = P(X \le g^{-1}(y)) \\ &= F_X(g^{-1}(y)). \end{aligned}$$

2.Suppose that g is a strictly decreasing function. For $y \in R_Y$,

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(g(X) \le y) = P(X \ge g^{-1}(y)) = 1 - P(X < g^{-1}(y)) \\ &= 1 - F_X(g^{-1}(y)). \end{aligned}$$

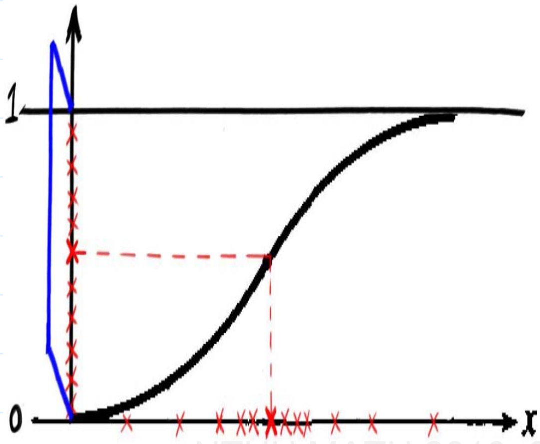
- Theorem. Let X be a continuous random variable whose cdf F_X possesses a unique inverse F_X^{-1}. Let $Z=F_X(X)$, then Z has a uniform distribution on [0, 1].

  Proof. For $0\leq z\leq 1$, $F_Z(z) = F_X(F_X^{-1}(z)) = z.$

- Theorem. Let U be a uniform random variable on [0, 1] and F is a cdf which possesses a unique inverse F^{-1}. Let $X=F^{-1}(U)$, then the cdf of X is F.

  Proof. $F_X(x) = F_U(F(x)) = P(U \leq F(x)) = F(x).$

- The 2 theorems are useful for pseudo-random number generation in computer simulation.



- X is r.v. $\Rightarrow$ $F(X)$ is r.v.
- $X_1, \ldots, X_n$: r.v.'s with cdf F $\Rightarrow$ $F(X_1), \ldots, F(X_n)$: r.v.'s with distribution Uniform(0, 1)
- $U_1, \ldots, U_n$: r.v.'s with distribution Uniform(0, 1) $\Rightarrow$ $F^{-1}(U_1), \ldots, F^{-1}(U_n)$: r.v.'s with cdf F

➢ The pdf of Y, denoted by f_Y

1. Suppose that g is a *differentiable* strictly increasing function. For $y\in R_Y$,

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(g^{-1}(y)) \\ &= f_X(g^{-1}(y))\frac{dg^{-1}(y)}{dy} = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|. \end{aligned}$$

2. Suppose that g is a *differentiable* strictly decreasing function. For $y\in R_Y$,

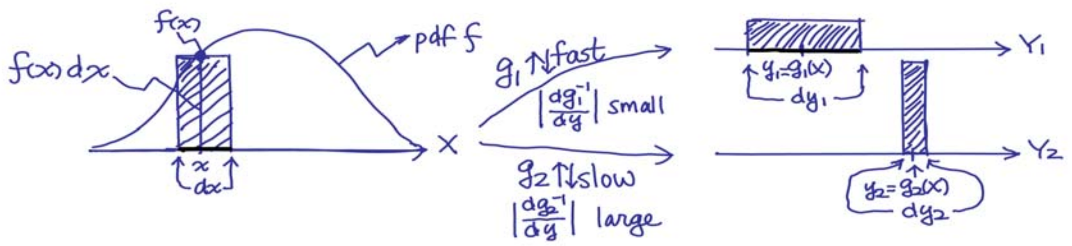
$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}(1-F_X(g^{-1}(y))) \\ &= -f_X(g^{-1}(y))\frac{dg^{-1}(y)}{dy} = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|. \end{aligned}$$

- Theorem. Let X be a continuous random variable with pdf f_X. Let $Y=g(X)$, where g is differentiable and strictly monotone. Then, the pdf of Y, denoted by f_Y, is

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|,$$

for y such that $y=g(x)$ for some x, and $f_Y(y)=0$ otherwise.

- **Q**: What is the role of $|dg^{-1}(y)/dy|$? How to interpret it?



➢Some Examples. Given the pdf f_X of random variable X,

- find the pdf f_Y of $Y=aX+b$, where $a\neq 0$.

$$y = g(x) = ax+b \;\Rightarrow\; x = g^{-1}(y) = \frac{y-b}{a} \;\Rightarrow\; \left|\frac{d}{dy}g^{-1}(y)\right| = \frac{1}{|a|}$$

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\cdot\frac{1}{|a|}$$

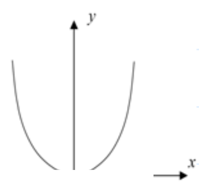
- find the pdf f_Y of $Y=1/X$.

$$y = g(x) = \frac{1}{x} \;\Rightarrow\; x = g^{-1}(y) = \frac{1}{y} \;\Rightarrow\; \left|\frac{d}{dy}g^{-1}(y)\right| = |-y^{-2}| = \frac{1}{y^2}$$

$$f_Y(y) = f_X\left(\frac{1}{y}\right)\cdot\frac{1}{y^2}$$

- find the cdf F_Y and pdf f_Y of $Y=X^2$.

  - $$\begin{aligned} F_Y(y) &= P(Y\le y) = P(-\sqrt{y}\le X\le\sqrt{y}) \\ &= P(X\in(-\infty,\sqrt{y}]) - P(X\in(-\infty,-\sqrt{y})) \\ &= \begin{cases} F_X(\sqrt{y}) - F_X(-\sqrt{y}), & \text{if } y>0, \\ 0, & \text{if } y\le 0. \end{cases} \end{aligned}$$



  - For $y>0$,

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right]$$

$$= f_X(\sqrt{y})\cdot\frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})\cdot\frac{-1}{2\sqrt{y}} = \frac{f_X(\sqrt{y})+f_X(-\sqrt{y})}{2\sqrt{y}}$$

For $y\le 0$, $f_Y(y)=0$.

• Expectation, Mean, and Variance

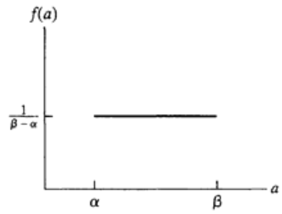
➢Definition. If X has a pdf f_X, then the *expectation* of X is defined by

$$E(X) = \int_{-\infty}^{\infty} x\cdot f_X(x)\,dx,$$

provided that the integral converges absolutely.

- Example (Uniform Distributions). If

$$f_X(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \text{if } \alpha < x \le \beta, \\ 0, & \text{otherwise,} \end{cases}$$



then

$$\begin{aligned} E(X) &= \int_\alpha^\beta x \cdot \frac{1}{\beta-\alpha}\, dx = \frac{1}{2}\cdot \frac{x^2}{\beta-\alpha}\bigg|_\alpha^\beta \\ &= \frac{1}{2}\cdot\frac{\beta^2-\alpha^2}{\beta-\alpha} = \frac{\alpha+\beta}{2}. \end{aligned}$$

➢Some properties of expectation

- Expectation of Transformation. If $Y=g(X)$, then

$$E(Y) = \int_{-\infty}^{\infty} y \cdot f_Y(y)\, dy = \int_{-\infty}^{\infty} g(x)\cdot f_X(x)\, dx,$$

provided that the integral converges absolutely.

proof. (homework)

- Expectation of Linear Function. $E(aX+b)=a\cdot E(X)+b$, since

$$\begin{aligned} E(aX+b) &= \int_{-\infty}^{\infty} (ax+b) f_X(x)\, dx \\ &= a\int_{-\infty}^{\infty} x\cdot f_X(x)\, dx + b\int_{-\infty}^{\infty} f_X(x)\, dx \\ &= a\cdot E(X) + b. \end{aligned}$$

➢Definition. If X has a pdf f_X, then the expectation of X is also called the *mean* of X or f_X and denoted by μ_X, so that

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x\cdot f_X(x)\, dx.$$

The *variance* of X is defined as

$$Var(X) = E[(X-\mu_x)^2] = \int_{-\infty}^{\infty} (x-\mu_X)^2 \cdot f_X(x)\, dx,$$

and denoted by σ_X^2. The σ_X is called the *standard deviation.*

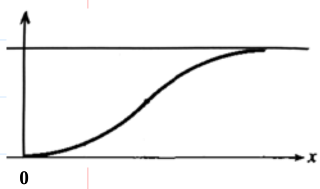
➢Some properties of mean and variance

- The mean and variance for continuous random variables have the same intuitive interpretation as in the discrete case.
- $Var(X) = E(X^2) - [E(X)]^2$
- Variance of Linear Function. $Var(aX+b)=a^2\cdot Var(X)$
- Theorem. For a nonnegative continuous random variable X,

$$E(X) = \int_0^\infty 1 - F_X(x)dx = \int_0^\infty P(X>x)dx.$$

Proof.

$$\begin{aligned} E(X) &= \int_0^\infty x\cdot f_X(x)\, dx \\ &= \int_0^\infty \left(\int_0^x 1\, dt\right) f_X(x)\, dx \\ &= \int_0^\infty \int_0^x f_X(x)\, dt\, dx \\ &= \int_0^\infty \int_t^\infty f_X(x)dx\, dt = \int_0^\infty 1 - F_X(t)dt. \end{aligned}$$



➢Example (Uniform Distributions)

$$E(X^2)=\int_{-\infty}^{\infty}\frac{x^2}{\beta-\alpha}dx=\frac{1}{3}\frac{x^3}{\beta-\alpha}\Big|_{\alpha}^{\beta}=\frac{\beta^3-\alpha^3}{3(\beta-\alpha)}=\frac{\beta^2+\alpha\beta+\alpha^2}{3}.$$

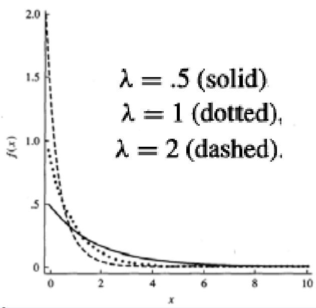
$$\begin{aligned} Var(X) &= E(X^2)-[E(X)]^2=\frac{\beta^2+\alpha\beta+\alpha^2}{3}-\left(\frac{\alpha+\beta}{2}\right)^2 \\ &= \frac{4(\beta^2+\alpha\beta+\alpha^2)-3(\beta^2+2\alpha\beta+\alpha^2)}{12}=\frac{(\beta-\alpha)^2}{12}. \end{aligned}$$

❖ **Reading**: textbook, Sec 5.1, 5.2, 5.3, 5.7

Some Common Continuous Distributions

- Uniform Distribution
  - ➢Summary for $X\sim$ Uniform(α, β)
    - Pdf: $$f(x)=\begin{cases} 1/(\beta-\alpha), & \text{if } \alpha<x\le\beta, \\ 0, & \text{otherwise}, \end{cases}$$
    - Cdf: $$F(x)=\begin{cases} 0, & \text{if } x\le\alpha, \\ (x-\alpha)/(\beta-\alpha), & \text{if } \alpha<x\le\beta, \\ 1, & \text{if } x>\beta. \end{cases}$$
    - Parameters: $-\infty<\alpha<\beta<\infty$
    - Mean: $E(X)=(\alpha+\beta)/2$
    - Variance: $Var(X)=(\beta-\alpha)^2/12$

- Exponential Distribution



➢For $\lambda>0$, the function

$$f(x)=\begin{cases} \lambda e^{-\lambda x}, & \text{if } x\ge 0, \\ 0, & \text{if } x<0, \end{cases}$$

is a pdf since (1) $f(x)\ge 0$ for all $x\in\mathbb{R}$, and (2)

$$\int_{-\infty}^{\infty} f(x)\,dx=\int_0^{\infty}\lambda e^{-\lambda x}\,dx=-e^{-\lambda x}\Big|_0^{\infty}=1.$$

- The distribution of a random variable X with this pdf is called the *exponential* distribution with parameter λ.

➢The cdf of an exponential r.v. is $F(x)=0$ for $x<0$, and for $x\ge 0$,

$$F(x)=P(X\le x)=\int_0^x\lambda e^{-\lambda y}\,dy=-e^{-\lambda y}\Big|_0^x=1-e^{-\lambda x}.$$

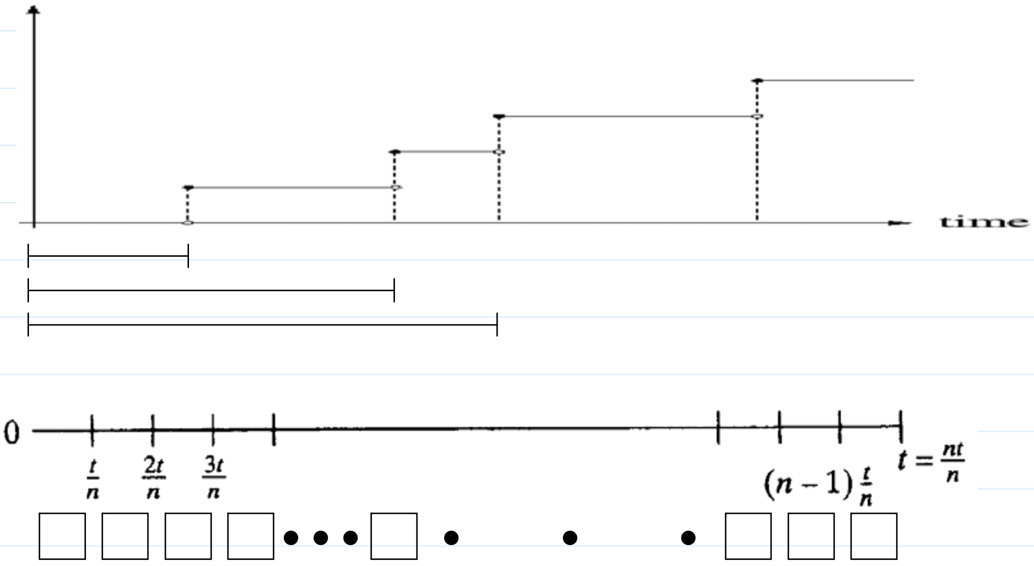
➢Theorem. The mean and variance of an exponential distribution with parameter λ are

$$\mu=1/\lambda \quad\text{and}\quad \sigma^2=1/\lambda^2.$$

Proof.
$$\begin{aligned} E(X) &= \int_0^{\infty}x\lambda e^{-\lambda x}dx=\int_0^{\infty}\frac{y}{\lambda}(\lambda e^{-y})\frac{1}{\lambda}dy \\ &= \frac{1}{\lambda}\int_0^{\infty}ye^{-y}dy=\frac{1}{\lambda}\Gamma(2)=\frac{1}{\lambda}. \\ E(X^2) &= \int_0^{\infty}x^2\lambda e^{-\lambda x}dx=\int_0^{\infty}\left(\frac{y}{\lambda}\right)^2(\lambda e^{-y})\frac{1}{\lambda}dy \\ &= \frac{1}{\lambda^2}\int_0^{\infty}y^2e^{-y}dy=\frac{1}{\lambda^2}\Gamma(3)=\frac{2}{\lambda^2}. \end{aligned}$$

➢Some properties

- The exponential distribution is often used to model the length of time until an event occurs
  - The parameter λ is called the rate and is the average number of events that occur in unit time. This gives an intuitive interpretation of $E(X)=1/\lambda$.



- Theorem (relationship between exponential, gamma, and Poisson distributions, Sec. 9.1). Let

  1. $T_1, T_2, T_3, \ldots,$ be independent and ~ exponential(λ),
  2. $S_k=T_1+\cdots+T_k$, $k=1, 2, 3, \ldots,$
  3. X_i be the number of S_k's that falls in the time interval $(t_{i-1}, t_i]$, $i=1, \ldots, m$.

  Then, (i) $X_1, \ldots, X_m$ are independent,

  (ii) $X_i \sim \text{Poisson}(\lambda(t_i - t_{i-1}))$,

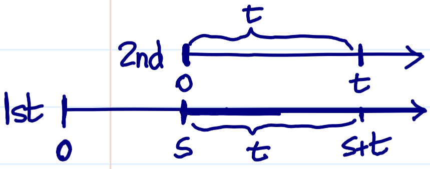
  (iii) $S_k \sim \text{Gamma}(k, \lambda)$.

  (iv) The reverse statement is also true.

  - The rate parameter λ is the same for both the Poisson and exponential random variable.
  - The exponential distribution can be thought of as the continuous analogue of the geometric distribution.

- Theorem. The exponential distribution (like the geometric distribution) is *memoryless*, i.e., for $s, t \geq 0$,

$$P(X > s + t | X > s) = P(X > t).$$



$P(X_1 > s+t | X_1 > s) = P(X_2 > t)$.

Proof.

$$P(X > s+t | X > s) = \frac{P(\{X > s+t\} \cap \{X > t\})}{P(\{X > s\})} = \frac{P(\{X > s+t\})}{P(\{X > s\})}$$

$$= \frac{1 - F_X(s+t)}{1 - F_X(s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t)$$

- This means that the distribution of the waiting time to the next event remains the same regardless of how long we have already been waiting.
- This only happens when events occur (or not) totally at random, i.e., independent of past history.
- Notice that it does not mean the two events $\{X > s+t\}$ and $\{X > t\}$ are independent.

➢ Summary for $X \sim$ Exponential(λ)

- Pdf: $$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$
- Cdf: $$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$
- Parameters: $\lambda > 0$.
- Mean: $E(X) = 1/\lambda$.
- Variance: $Var(X) = 1/\lambda^2$.

• Gamma Distribution

➢ Gamma Function

- Definition. For $\alpha > 0$, the *gamma function* is defined as
$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx.$$
- $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$ (exercise)
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$
  Proof. By integration by parts,
$$\begin{aligned} \Gamma(\alpha+1) &= \int_0^\infty x^\alpha e^{-x}\, dx \\ &= -x^\alpha e^{-x}\Big|_0^\infty + \int_0^\infty \alpha x^{\alpha-1} e^{-x}\, dx = \alpha\Gamma(\alpha). \end{aligned}$$
- $\Gamma(\alpha) = (\alpha-1)!$ if α is an integer
$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) = \cdots = (\alpha-1)(\alpha-2)\cdots\Gamma(1) = (\alpha-1)!$$
- $\Gamma(\alpha/2) = \frac{\sqrt{\pi}(\alpha-1)!}{2^{\alpha-1}(\frac{\alpha-1}{2})!}$ if α is an odd integer
$$\Gamma(\tfrac{\alpha}{2}) = (\tfrac{\alpha-2}{2})\Gamma(\tfrac{\alpha}{2}-1) = \cdots = (\tfrac{\alpha-2}{2})(\tfrac{\alpha-4}{2})\cdots\tfrac{1}{2}\Gamma(\tfrac{1}{2})$$
- Gamma function is a generalization of the factorial functions

➢ For α, $\lambda > 0$, the function
$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases}$$

is a pdf since (1) $f(x) \geq 0$ for all $x \in \mathrm{R}$, and (2)

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} y^{\alpha-1} e^{-y}\,dy = 1.$$

- The distribution of a random variable X with this pdf is called the *gamma* distribution with parameters α and λ.

➢ The cdf of gamma distribution can be expressed in terms of the *incomplete gamma function*, i.e., $F(x)$=0 for x<0, and for $x \geq 0$,

$$F(x) = \int_0^x \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy = \frac{1}{\Gamma(\alpha)} \int_0^{\lambda x} z^{\alpha-1} e^{-z} dz \equiv \frac{\gamma(\alpha, \lambda x)}{\Gamma(\alpha)}.$$

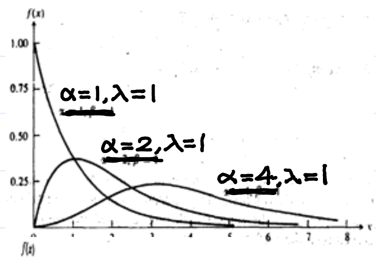
➢ Theorem. The mean and variance of a gamma distribution with parameter α and λ are

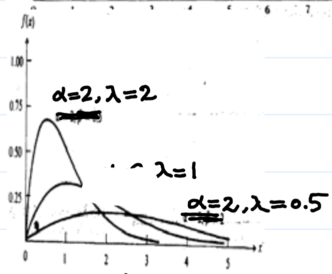
$$\mu = \alpha/\lambda \quad \text{and} \quad \sigma^2 = \alpha/\lambda^2.$$

Proof. $$E(X) = \int_0^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} \int_0^{\infty} \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{(\alpha+1)-1} e^{-\lambda x}\,dx = \frac{\alpha}{\lambda}.$$

$$E(X^2) = \int_0^{\infty} x^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}} \int_0^{\infty} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{(\alpha+2)-1} e^{-\lambda x}\,dx = \frac{\alpha(\alpha+1)}{\lambda^2}.$$

- (exercise) $E(X^k) = \frac{\Gamma(\alpha+k)}{\lambda^k \Gamma(\alpha)}$, for $0 < k$, and $E(\frac{1}{X^k}) = \frac{\lambda^k \Gamma(\alpha-k)}{\Gamma(\alpha)}$, for $0 < k < \alpha$.





➢ Some properties

- The gamma distribution can be used to model the waiting time until a number of random events occurs
  - When α=1, it is exponential(λ)
  - $T_1, \ldots, T_n$: independent exponential(λ) r.v.'s $\Rightarrow T_1 + \cdots + T_n \sim$ Gamma(n, λ)
  - Gamma distribution can be thought of as a continuous analogue of the negative binomial distribution
    - A summary

	Discrete Time Version	Continuous Time Version
number of events	binomial	Poisson
waiting time until 1 event occurs	geometric	exponential
waiting time until r events occur	negative binomial	gamma

- α is called shape parameter and λ scale parameter (**Q**: how to interpret α and λ from the viewpoint of waiting time?)

- A special case of the gamma distribution occurs when $\alpha=n/2$ and $\lambda=1/2$ for some positive integer n. This is known as the Chi-squared distribution with n degrees of freedom (Chapter 6)

➢ Summary for X ~ Gamma(α, λ)

- Pdf: $$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$
- Cdf: $F(x) = \gamma(\alpha, \lambda x)/\Gamma(\alpha)$.
- Parameters: $\alpha, \lambda > 0$.
- Mean: $E(X) = \alpha/\lambda$.
- Variance: $Var(X) = \alpha/\lambda^2$.

• Beta Distribution

➢ Beta Function: $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

➢ For $\alpha, \beta > 0$, the function

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{if } 0 \leq x \leq 1, \\ 0, & \text{otherwise}, \end{cases}$$

is a pdf (exercise).

- The distribution of a random variable X with this pdf is called the *beta* distribution with parameters α and β.

➢ The cdf of beta distribution can be expressed in terms of the *incomplete beta function*, i.e., $F(x)=0$ for $x<0$, $F(x)=1$ for $x>1$, and for $0 \leq x \leq 1$,

$$F(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x y^{\alpha-1}(1-y)^{\beta-1}\, dy \equiv \frac{B(x;\alpha,\beta)}{B(\alpha,\beta)}.$$

$$= \text{(exercise)} \sum_{i=\alpha}^{\alpha+\beta-1} \frac{(\alpha+\beta-1)!}{i!(\alpha+\beta-1-i)!} x^i (1-x)^{\alpha+\beta-1-i},$$

for *integer values* of α and β

➢ Theorem. The mean and variance of a beta distribution with parameters α and β are

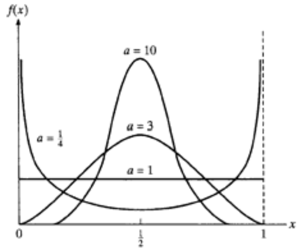
$$\mu = \frac{\alpha}{\alpha+\beta} \quad \text{and} \quad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Proof.

$$E(X) = \int_0^\infty x \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx$$
$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} \int_0^\infty \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta)} x^{(\alpha+1)-1}(1-x)^{\beta-1}\, dx$$
$$= \frac{\alpha}{\alpha+\beta}.$$

$$E(X^2) = \int_0^\infty x^2 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx$$
$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)} \int_0^\infty \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+2)\Gamma(\beta)} x^{(\alpha+2)-1}(1-x)^{\beta-1}\,dx$$
$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.$$

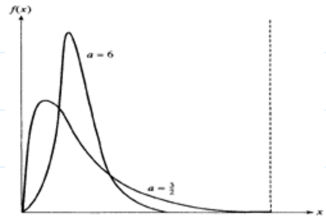
➢ Some properties



- When α=β=1, the beta distribution is the same as the uniform(0, 1).
- Whenever α=β, the beta distribution is symmetric about x=0.5, i.e.,

  f(0.5–Δ)=f(0.5+Δ).

  - As the common value of α and β increases, the distribution becomes more peaked at x=0.5 and there is less probability outside of the central portion.



- When β>α, values close to 0 become more likely than those close to 1; when β<α, values close to 1 are more likely than those close to 0 (**Q**: How to connect it with $E(X)$?)

➢ Summary for X ~ Beta(α, β)

- Pdf: $$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{if } 0 \le x \le 1, \\ 0, & \text{otherwise,} \end{cases}$$
- Cdf: $F(x) = B(x;\alpha,\beta)/B(\alpha,\beta)$.
- Parameters: α, β > 0.
- Mean: $E(X)$ = α/(α + β).
- Variance: $Var(X)$ = [α(α + 1)]/[(α + β)(α + β +1)].

• Normal Distribution

➢ For μ∈ℝ and σ>0, the function

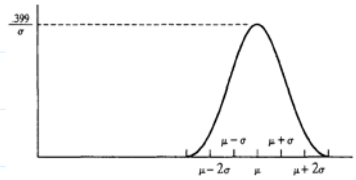
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$$

is a pdf since (1) $f(x) \ge 0$ for all $x \in \mathbb{R}$, and (2)

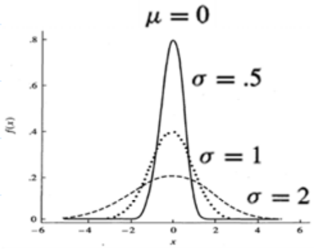
$$\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{y^2}{2}}\,dy \equiv \frac{I}{\sqrt{2\pi}},$$

and $$I^2 = \left(\int_{-\infty}^\infty e^{-\frac{x^2}{2}}\,dx\right)\left(\int_{-\infty}^\infty e^{-\frac{y^2}{2}}\,dy\right)$$
$$= \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-\frac{x^2+y^2}{2}}\,dxdy = \int_0^\infty \int_0^{2\pi} e^{-\frac{r^2}{2}} r\,d\theta dr$$
$$= 2\pi \int_0^\infty re^{-\frac{r^2}{2}}\,dr = -2\pi e^{-\frac{r^2}{2}}\Big|_0^\infty = 2\pi.$$

- The distribution of a random variable X with this pdf is called the *normal* (*Gaussian*) distribution with parameters μ and σ, denoted by $N(\mu, \sigma^2)$.



- The normal pdf is a bell-shaped curve.
  - It is symmetric about the point μ, i.e., $f(\mu+\Delta)=f(\mu-\Delta)$ and falls off in the rate determined by σ.
  - The pdf has a maximum at μ (can be shown by differentiation) and the maximum height is $1/(\sigma\sqrt{2\pi})$.



➢The cdf of normal distribution does not have a close form.

➢Theorem. The mean and variance of a $N(\mu, \sigma^2)$ distribution are μ and σ^2, respectively.

- μ: location parameter; σ^2: scale parameter

Proof. $$E(X) = \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \int_{-\infty}^{\infty}(\sigma y+\mu)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,dy$$
$$= \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2}}\,dy + \mu\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,dy$$
$$= \frac{\sigma}{\sqrt{2\pi}}\cdot 0 + \mu\cdot 1 = \mu.$$

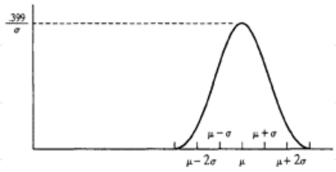
$$E(X^2) = \int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \int_{-\infty}^{\infty}(\sigma y+\mu)^2\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,dy$$
$$= \sigma^2\int_{-\infty}^{\infty} y^2\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,dy + \frac{2\mu\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2}}\,dy$$
$$+\mu^2\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\,dy$$
$$= \sigma^2\cdot 1 + \frac{2\mu\sigma}{\sqrt{2\pi}}\cdot 0 + \mu^2\cdot 1 = \sigma^2+\mu^2.$$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y\left(ye^{-\frac{y^2}{2}}\right)dy = \frac{1}{\sqrt{2\pi}}y\left(-e^{-y^2/2}\right)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\left(-e^{-\frac{y^2}{2}}\right)dy$$

➢Some properties

- Normal distribution is one of the most widely used distribution. It can be used to model the distribution of many natural phenomena.
- Theorem. Suppose that $X\sim N(\mu, \sigma^2)$. The random variable $Y=aX+b$, where $a\neq 0$, is also normally distributed with parameters $a\mu+b$ and $a^2\sigma^2$, i.e., $Y\sim N(a\mu+b, a^2\sigma^2)$.

  Proof. $$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|} = \frac{1}{\sqrt{2\pi}|a|\sigma}e^{-\frac{[y-(a\mu+b)]^2}{2\sigma^2a^2}}.$$



- Corollary. If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X-\mu}{\sigma}$$

is a normal random variable with parameters 0 and 1, i.e., $N(0, 1)$, which is called *standard normal distribution*.

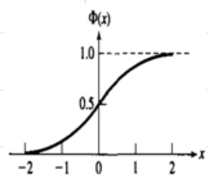
- The $N(0, 1)$ distribution is very important since properties of any other normal distributions can be found from those of the standard normal.
  - *The cdf of* $N(0, 1)$ *is usually denoted by* Φ.
  - Theorem. Suppose that $X \sim N(\mu, \sigma^2)$. The cdf of X is

$$F_X(x) = \Phi\left(\tfrac{x-\mu}{\sigma}\right).$$

Proof. $F_X(x) = F_Z\left(\tfrac{x-\mu}{\sigma}\right) = \Phi\left(\tfrac{x-\mu}{\sigma}\right).$

  - Example. Suppose that $X \sim N(\mu, \sigma^2)$. For $-\infty < a < b < \infty$,

$$\begin{aligned} P(a < X < b) &= P\left(\tfrac{a-\mu}{\sigma} < \tfrac{X-\mu}{\sigma} < \tfrac{b-\mu}{\sigma}\right) \\ &= P\left(\tfrac{a-\mu}{\sigma} < Z < \tfrac{b-\mu}{\sigma}\right) = P\left(Z < \tfrac{b-\mu}{\sigma}\right) - P\left(Z < \tfrac{a-\mu}{\sigma}\right) \\ &= \Phi\left(\tfrac{b-\mu}{\sigma}\right) - \Phi\left(\tfrac{a-\mu}{\sigma}\right). \end{aligned}$$



  - Table 5.1 (textbook, p.201) gives values of Φ. To read the table:
    1. Find the first value of z up to the first place of decimal in the left hand column.
    2. Find the second place of decimal across the top row.
    3. The value of $\Phi(z)$ is where the row from the first step and the column from the second step intersect.

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF x

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
					• • •					
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

    - For the values greater than z=3.49, $\Phi(z) \approx 1$.
    - For negative values of z, use $\Phi(-z)=1-\Phi(z)$

- Normal distribution plays a central role in the limit theorems of probability (e.g., CLT, chapter 8)

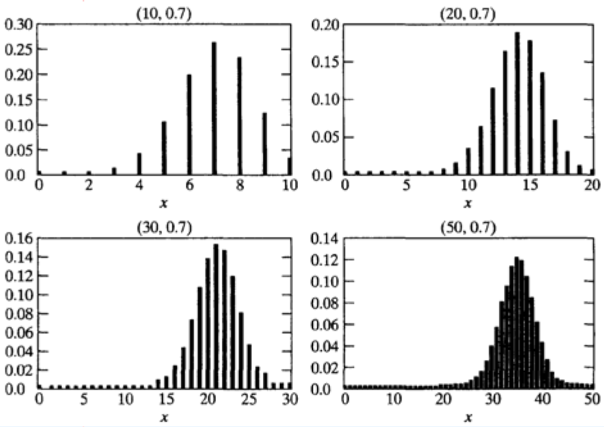
➢Normal approximation to the Binomial

- <u>Recall</u>. Poisson approximation to Binomial
- Theorem. Suppose that X_n~binomial(n, p). Define

$$Z_n = (X_n - np)/\sqrt{np(1-p)}.$$

Then, as $n \to \infty$, the distribution of Z_n converge to the $N(0, 1)$ distribution, i.e., $F_{Z_n}(z) = P(Z_n \le z) \to \Phi(z)$.

Proof. It is a special case of the CLT in Chapter 8.



- Plot the pmf of X_n~binomial(n, p)
- Superimpose the pdf of $Y_n \sim N(\mu_n, \sigma_n^2)$ with $\mu_n = np$ and $\sigma_n^2 = np(1-p)$.
- When n is sufficiently large, the normal pdf approximates the binomial pmf.
- $Z_n \overset{d}{\approx} (Y_n - \mu_n)/\sigma_n$
- The size of n to achieve a good approximation depends on the value of p.
  - For p near 0.5 ⇒ moderate n is enough
  - For p close to zero or one ⇒ require much larger n

- Continuity Correction
  - **Q**: Why need continuity correction? <u>Ans</u>. The binomial(n, p) is a discrete random variable and we are approximating it with a continuous random variable.
    - For example, suppose X~binomial(50, 0.4) and we want to find $P(X=18)$, which is larger than 0.
    - With the normal pdf, however, $P(Y=18)=0$ since we are using a continuous distribution
  - Instead, we make a continuity correction,

$$
\begin{aligned}
P(X=18) &= P(17.5 < X < 18.5) = P\left(\tfrac{17.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}} < Z_n < \tfrac{18.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}}\right) \\
&\approx P\left(\tfrac{17.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}} < Z < \tfrac{18.5-(50\cdot 0.4)}{\sqrt{50\cdot 0.4\cdot 0.6}}\right) \\
&= P\left(-\tfrac{2.5}{\sqrt{12}} < Z < -\tfrac{1.5}{\sqrt{12}}\right) = P\left(Z < -\tfrac{1.5}{\sqrt{12}}\right) - P\left(Z < -\tfrac{2.5}{\sqrt{12}}\right) \\
&= \Phi\left(-\tfrac{1.5}{\sqrt{12}}\right) - \Phi\left(-\tfrac{2.5}{\sqrt{12}}\right) = \left(1 - \Phi\left(\tfrac{1.5}{\sqrt{12}}\right)\right) - \left(1 - \Phi\left(\tfrac{2.5}{\sqrt{12}}\right)\right) \\
&= \Phi\left(2.5/\sqrt{12}\right) - \Phi\left(1.5/\sqrt{12}\right).
\end{aligned}
$$

and can obtain the approximate value from Table 5.1.

  - Similary, $P(X \ge 30) = P(X > 29.5) = P\left(Z_n > \frac{29.5-(50\cdot 0.4)}{\sqrt{12}}\right)$

and $\approx P(Z > 9.5/\sqrt{12}) = 1 - \Phi(9.5/\sqrt{12}).$

$$
\begin{aligned}
P(10 \le X \le 30) &= P(9.5 < X < 30.5) \\
&= P\left(\frac{9.5-(50\cdot 0.4)}{\sqrt{12}} < Z_n < \frac{30.5-(50\cdot 0.4)}{\sqrt{12}}\right) \\
&\approx P(-10.5/\sqrt{12} < Z < 10.5/\sqrt{12}) \\
&= \Phi(10.5/\sqrt{12}) - \Phi(-10.5/\sqrt{12}) = 2\cdot\Phi(10.5/\sqrt{12}) - 1.
\end{aligned}
$$

- Summary for $X \sim$ Normal(μ, σ^2)
  - Pdf: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$
  - Cdf: no close form, but usually denoted by $\Phi((x-\mu)/\sigma)$.
  - Parameters: $\mu \in \mathbb{R}$ and $\sigma > 0$.
  - Mean: $E(X) = \mu$.
  - Variance: $Var(X) = \sigma^2$.

- Weibull Distribution
  - For $\alpha, \beta > 0$ and $\nu \in \mathbb{R}$, the function

$$
f(x) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta}, & \text{if } x \ge \nu, \\ 0, & \text{if } x < \nu, \end{cases}
$$

is a pdf since (1) $f(x) \ge 0$ for all $x \in \mathbb{R}$, and (2)

$$
\begin{aligned}
\int_{-\infty}^{\infty} f(x)\, dx &= \int_{\nu}^{\infty} \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta} dx \\
&= \int_0^\infty e^{-y}\, dy = -e^{-y}\Big|_0^\infty = 1.
\end{aligned}
$$

  - The distribution of a random variable X with this pdf is called the *Weibull* distribution with parameters α, β, and ν .

- (exercise) The cdf of Weibull distribution is

$$
F(x) = \begin{cases} 1 - e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta}, & \text{if } x \ge \nu, \\ 0, & \text{if } x < \nu. \end{cases}
$$

- Theorem. The mean and variance of a Weibull distribution with parameter α, β, and ν are

$$
\mu = \alpha\Gamma\left(1 + \frac{1}{\beta}\right) + \nu \quad \text{and}
$$

$$
\sigma^2 = \alpha^2\left\{\Gamma\left(1+\frac{2}{\beta}\right) - \left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right\}.
$$

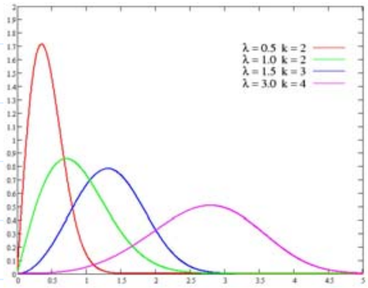
Proof.

$$
\begin{aligned}
E(X) &= \int_v^\infty x \cdot \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta} dx \\
&= \int_0^\infty (\alpha y^{1/\beta} + \mu) e^{-y}\, dy \\
&= \alpha \int_0^\infty y^{1/\beta} e^{-y}\, dy + \mu \int_0^\infty e^{-y}\, dy = \alpha\Gamma\left(\frac{1}{\beta}+1\right) + \mu
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \int_v^\infty x^2 \cdot \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\nu}{\alpha}\right)^\beta} dx \\
&= \int_0^\infty (\alpha y^{1/\beta} + \mu)^2 e^{-y}\, dy \\
&= \alpha^2 \int_0^\infty y^{2/\beta} e^{-y}\, dy + 2\alpha\mu \int_0^\infty y^{1/\beta} e^{-y}\, dy + \mu^2 \int_0^\infty e^{-y}\, dy \\
&= \alpha^2\Gamma\left(\frac{2}{\beta}+1\right) + 2\alpha\mu\Gamma\left(\frac{1}{\beta}+1\right) + \mu^2
\end{aligned}
$$

➢ Some properties

- Weibull distribution is widely used to model lifetime.
- α: scale parameter; β: shape parameter; ν: location parameter
- Theorem. If X~exponential(λ), then



$$Y = \alpha\,(\lambda X)^{1/\beta} + \mu$$

is distributed as Weibull with parameter α, β, and ν (exercise).

• Cauchy Distribution

➢ For μ∈ℝ and σ>0, the function

$$f(x) = \tfrac{\sigma}{\pi}\tfrac{1}{\sigma^2+(x-\mu)^2}, \quad -\infty < x < \infty,$$

is a pdf since (1) $f(x) \geq 0$ for all $x \in \mathbb{R}$, and (2)

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \int_{-\infty}^{\infty} \tfrac{\sigma}{\pi}\tfrac{1}{\sigma^2+(x-\mu)^2}\,dx \\ &= \int_{-\infty}^{\infty} \tfrac{1}{\pi}\tfrac{1}{1+y^2}\,dy = \tfrac{1}{\pi}\tan^{-1}(y)\Big|_{-\infty}^{\infty} = 1. \end{aligned}$$

- The distribution of a random variable X with this pdf is called the *Cauchy* distribution with parameters μ and σ, denoted by Cauchy(μ, σ).

➢ The cdf of Cauchy distribution is

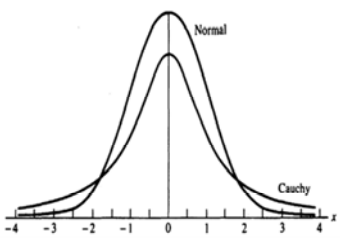
$$F(x) = \int_{-\infty}^{x} \tfrac{\sigma}{\pi}\tfrac{1}{\sigma^2+(y-\mu)^2}\,dy = \tfrac{1}{2} + \tfrac{1}{\pi}\tan^{-1}\left(\tfrac{x-\mu}{\sigma}\right)$$

for $-\infty<x<\infty$. (exercise)

➢ The mean and variance of Cauchy distribution do not exist because the integral does not converge absolutely

➢ Some properties

- Cauchy is a heavy tail distribution
- μ: location parameter; σ: scale parameter
- Theorem. If X~Cauchy(μ, σ), then $aX+b$~Cauchy(aμ+b, |a|σ). (exercise)



❖ **Reading**: textbook, Sec 5.4, 5.5, 5.6