Chapter 7

Special Continuous Distributions

Outline

- 7.1 Uniform random variable
- 7.2 Normal random variable
- 7.3 Exponential random variable
- 7.4 Gamma distribution
- 7.5 Beta distribution
- 7.6 Survival analysis and hazard function

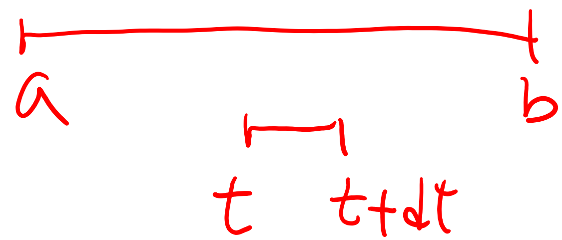
Section 7.1

Uniform Random Variables

Definition of Uniform Random Variables

- **Definition 7.1** A random variable X is said to be uniformly distributed over an interval (a, b) if its probability density function is given by (7.1).

$$F(t) = \begin{cases} 0 & t < a \\ \dfrac{t-a}{b-a} & a \le t < b \\ 1 & t \ge b \end{cases}$$



$$f(t) = F'(t) = \begin{cases} \dfrac{1}{b-a} & \text{if } a < t < b \\ 0 & \text{otherwise} \end{cases} \qquad (7.1)$$

An Alternative Definition

- Another way of reaching this definition is to note that $f(x)$ is a measure that determines how likely it is for X to be close to x.
- Since for all $x \in (a, b)$ the probability that X is close to x is the same, f should be a nonzero constant on (a, b); zero, elsewhere.
- Therefore,

$$f(x) = \begin{cases} c & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

- Since f is a probability density function, its integral over (a, b) is one.
- Thus,

$$c = \frac{1}{b-a}.$$

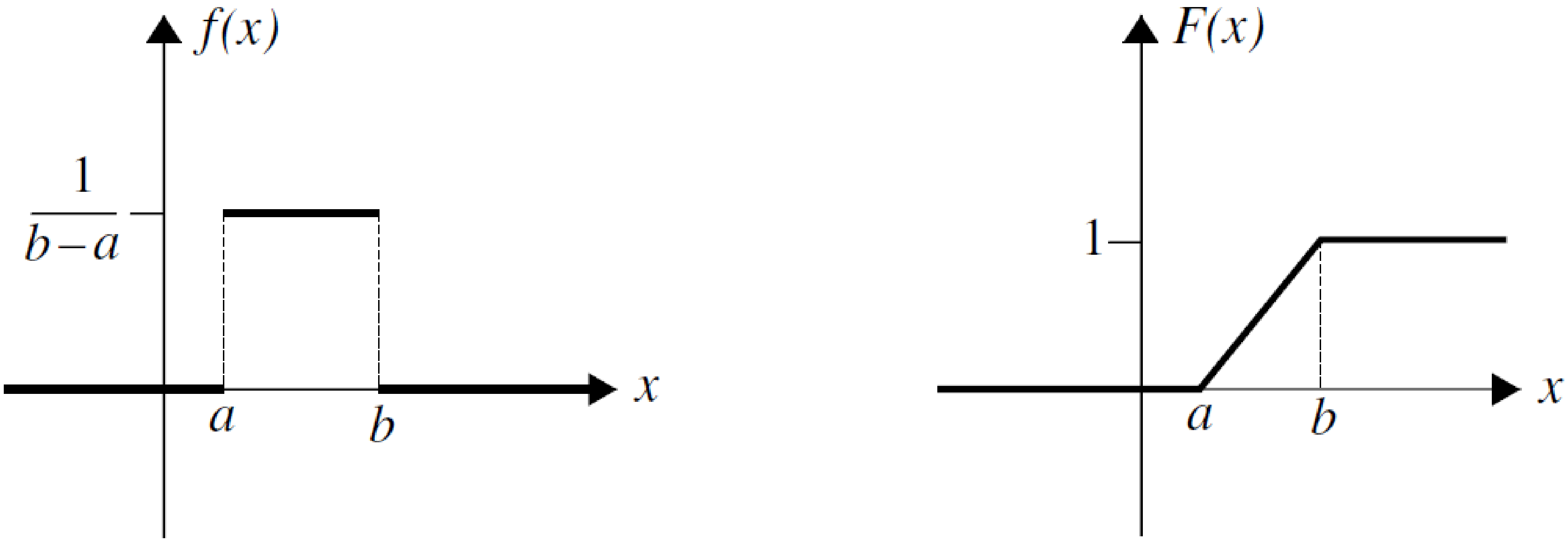


Figure 7.1 Density and distribution functions of a uniform random variable.

The Expectation and Variance

If X is uniformly distributed over an interval (a, b), then

$$E(X) = \frac{a+b}{2}, \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}, \qquad \sigma_X = \frac{b-a}{\sqrt{12}}.$$

$$E(X) = \int_a^b x\,\frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b = \frac{1}{b-a}\left(\frac{1}{2}b^2 - \frac{1}{2}a^2\right)$$

$$= \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}.$$

To find $\mathrm{Var}(X)$, note that

$$E(X^2) = \int_a^b x^2\frac{1}{b-a}dx = \frac{1}{3}\frac{b^3-a^3}{b-a} = \frac{1}{3}(a^2+ab+b^2).$$

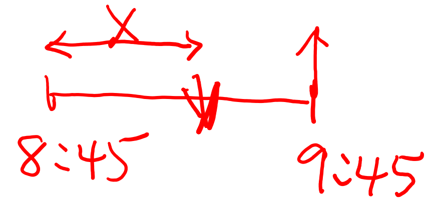
Hence

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{1}{3}(a^2+ab+b^2) - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

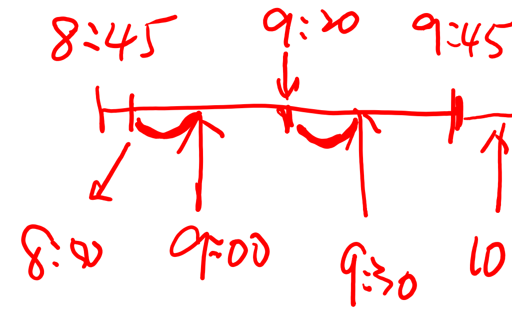
Example 7.1

- Starting at 5:00 A.M., every half hour there is a flight from San Francisco airport to Los Angeles International airport.
- Suppose that none of these planes is completely sold out and that they always have room for passengers.
- A person who wants to fly to L.A. arrives at the airport at a random time between 8:45 A.M. and 9:45 A.M.
- Find the probability that she waits
  - (a) at most 10 minutes;
  - (b) at least 15 minutes.

Solution of Example 7.1



- Let the passenger arrive at the airport X minutes past 8:45.
- Then X is a uniform random variable over the interval $(0, 60)$ with density function



$$f(x) = \begin{cases} 1/60 & \text{if } 0 < x < 60 \\ 0 & \text{otherwise.} \end{cases}$$

- Now the passenger waits at most 10 minutes if she arrives between 8:50 and 9:00 or 9:20 and 9:30; that is, if $5 < X < 15$ or $35 < X < 45$.

$P(X \in (5,15)$ or $X \in (35,45))$

- Solution to (a) is

$$P(5 < X < 15) + P(35 < X < 45) = \int_5^{15} \frac{1}{60} dx + \int_{35}^{45} \frac{1}{60} dx = \frac{1}{3}.$$

- The passenger waits at least 15 minutes if she arrives between 9:00 and 9:15 or 9:30 and 9:45; that is, if $15 < X < 30$ or $45 < X < 60$.
- Thus the answer to (b) is

$$P(15 < X < 30) + P(45 < X < 60) = \int_{15}^{30} \frac{1}{60}\,dx + \int_{45}^{60} \frac{1}{60}\,dx = \frac{1}{2}.$$

Example 7.3 Bertrand's Paradox

- What is the probability that a random chord of a circle is longer than a side of an equilateral triangle inscribed into the circle?
  - 1/2 ?
  - 1/3 ?
  - 1/4 ?

- The first interpretation of "random chords"

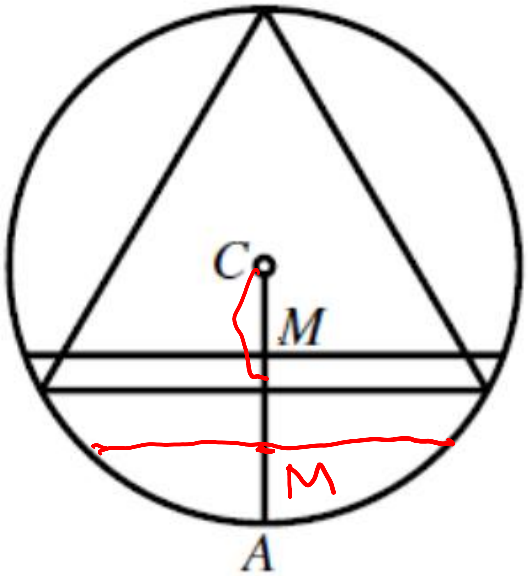


Figure 7.2 First interpretation.

- The second interpretation of “random chords”

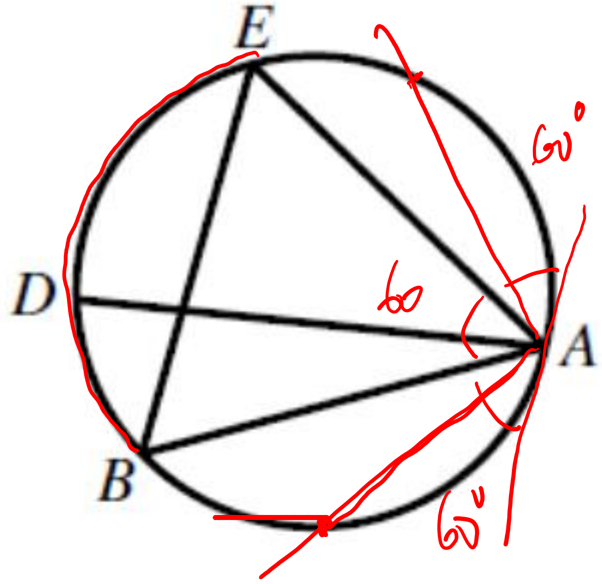
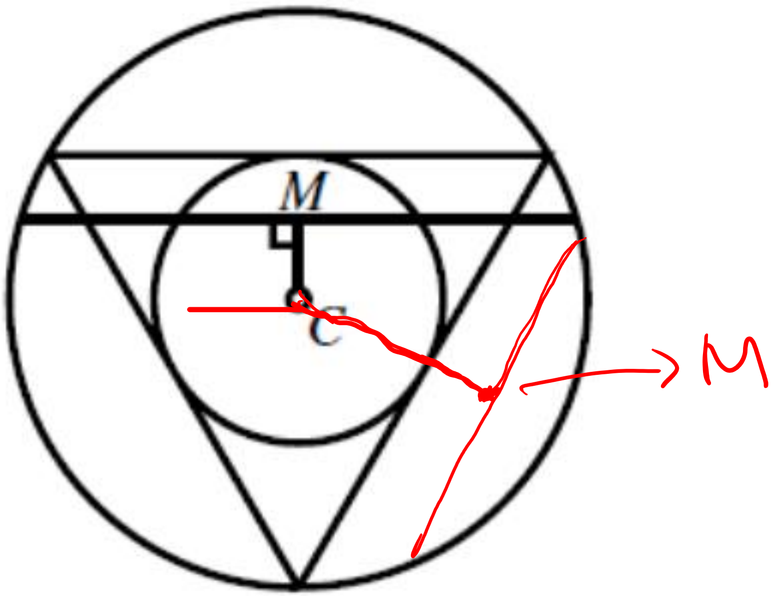


Figure 7.3 Second interpretation.

$$\frac{60^\circ}{180^\circ} = \frac{1}{3}$$

- The third interpretation of random chords



$$\frac{\pi(r/2)^2}{\pi r^2} = \frac{1}{4}.$$

Figure 7.4 Third interpretation.

- Three different interpretations of random chord resulted in three different answers.
- Because of this, Bertrand's problem was once considered a paradox.
  - Bertrand seriously doubted that probability could be defined on infinite sample spaces.
- At that time, one did not pay attention to the fact that the three interpretations correspond to three different experiments concerning the selection of a random chord.
- In this process we are dealing with three different probability functions defined on the same set of events.

Midterm exam 2

- Chapters 4, 5 (not including hypergeometric random variables), and 6
- 7.1

- You are required to bring your scientific calculator.
- Selected formula will be given.

Homework 8

- Section 6.1: 9, 12, 13
- Section 6.2: 6, 7, 9
- Section 6.3: 10, 14, 18, 19
- Section 7.1: 9, 12, 15, 17

- No need to submit your answers.
- Solution will be distributed in the weekend.

Section 7.2

Normal Random Variables

Limit of Binomial Distributions

- Poisson proved that binomial distributions converge to Poisson distributions as n becomes large
  - $n \to \infty$
  - $p \to 0$
  - $np = \lambda$

De Moivre's Theorem (1718)

- Let X be a binomial random variable with parameters n and ½.
- Then for any numbers a and $b, a < b$,

$$\lim_{n\to\infty} P\left(a < \frac{X - \left(\frac{1}{2}\right)n}{\left(\frac{1}{2}\right)\sqrt{n}} < b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-t^2/2}dt$$

- Note that in this formula $\left(\frac{1}{2}\right)n = E(X)$ and $\left(\frac{1}{2}\right)\sqrt{n} = \sigma_X$.

$X \sim \text{binomial}(n, p)$

$E(X) = np$

$Var(X) = np(1-p)$

De Moivre-Laplace Theorem (1812)

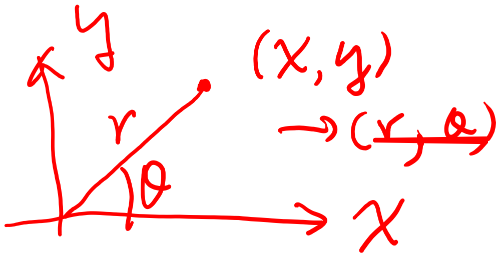
- Let X be a binomial random variable with parameters n and p.
- Then for any numbers a and b, $a < b$,

$$\lim_{n\to\infty} P\left(a < \frac{X - np}{\sqrt{np(1-p)}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

- Note that $np = E(X)$ and $\sqrt{np(1-p)} = \sigma_X$.

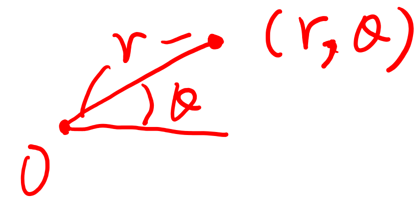
Gauss's Technique

- Gauss showed that $\Phi(t)$ is a distribution function, where



$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx$$

- Define



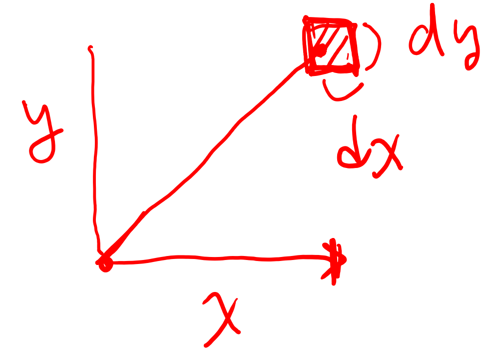
$$I = \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

- To show that $\Phi(\infty) = 1$, Gauss showed that $I^2 = 2\pi$.

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx\right)\left(\int_{-\infty}^{\infty} e^{-x^2/2} dx\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy$$

- We change the variables to polar coordinates.
- Let $x = r\cos\theta$ and $y = r\sin\theta$.
- We get

$$dx\,dy = r\,d\theta\,dr$$



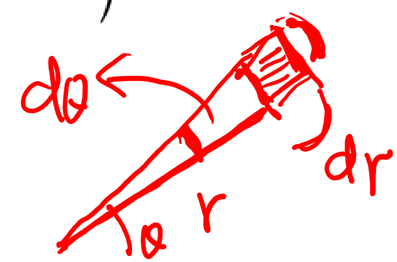
area = dxdy

- Thus,

$$I^2 = \int_0^\infty \int_0^{2\pi} e^{-r^2/2} r\,d\theta\,dr = \int_0^\infty e^{-r^2/2} r \left(\int_0^{2\pi} d\theta\right) dr$$

$$= 2\pi \int_0^\infty r e^{-r^2/2}\,dr = 2\pi \left[-e^{-r^2/2}\right]_0^\infty = 2\pi.$$

rdθ



- Thus, $I = \sqrt{2\pi}$ and $\Phi(\infty) = 1$.
- $\Phi(t)$ is a distribution function, since it is increasing, continuous and $\Phi(\infty) = 1$.

Definition of Standard Normal Distributions

- **Definition.** A random variable X is called standard normal if its distribution function is

$$P(X \leq t) = \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\, dx.$$

- By the fundamental theorem of calculus, f, the density function of a standard normal random variable, is given by

$$f(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Graphical Illustration

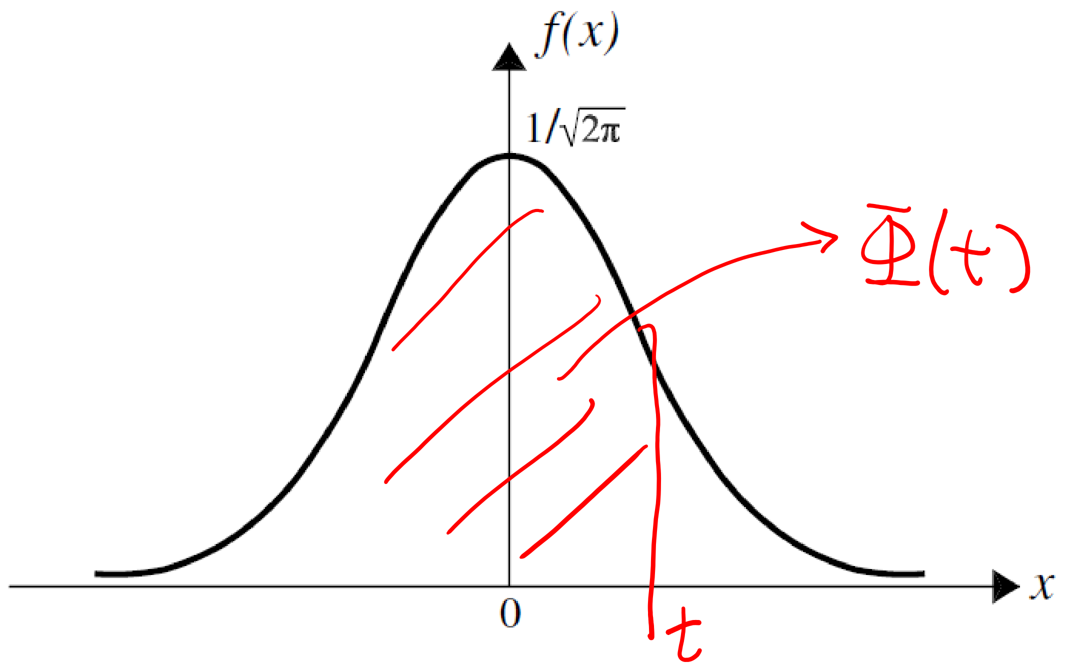


Figure 7.5 Graph of the standard normal density function.

Density Function of Standard Normal Random Variables

$$f(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

A Simple Property

- $\Phi(t)$ is the area under the (pdf) probability density function of the standard normal random variable from $-\infty$ to t.
- Since $\Phi(\infty) = 1$ and the pdf is symmetric about the y-axis,

$$\Phi(0) = \frac{1}{2} \qquad \text{and} \qquad \Phi(-t) = 1 - \Phi(t).$$

- To see this,

$u = -x$

$$\Phi(-t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-t} e^{-x^2/2}\, dx$$

$$\Phi(-t) = -\frac{1}{\sqrt{2\pi}} \int_{\infty}^{t} e^{-u^2/2}\, du = \frac{1}{\sqrt{2\pi}} \int_{t}^{\infty} e^{-u^2/2}\, du$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2}\, du - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2}\, du$$

$$= 1 - \Phi(t).$$

Correction for Continuity

- The De Moivre-Laplace theorem approximates a discrete distribution by a continuous distribution.
- One needs to make correction for continuity.

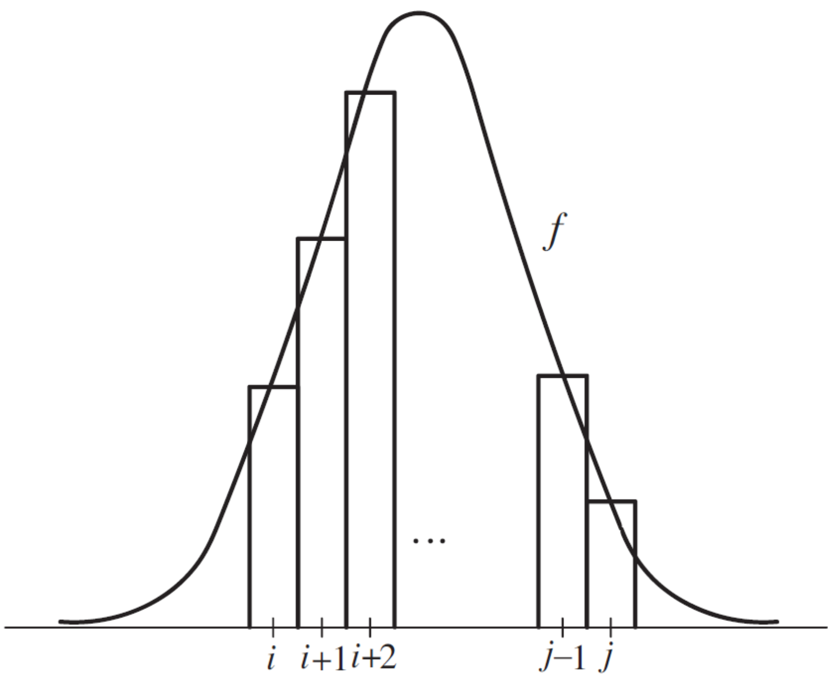


Figure 7.7 Histogram of X and the probability density function f.

Correction for Continuity

- $P(i \leq X \leq j) \approx \int_{i-1/2}^{j+1/2} f(x)\,dx$
- $P(X = k) \approx \int_{k-1/2}^{k+1/2} f(x)\,dx$
- $P(X \geq i) \approx \int_{i-1/2}^{\infty} f(x)\,dx$
- $P(X \leq j) \approx \int_{-\infty}^{j+1/2} f(x)\,dx$

Example 7.4

- Suppose that of all the clouds that are seeded with silver iodide, 58% show splendid growth.
- If 60 clouds are seeded with silver iodide, what is the probability that exactly 35 show splendid growth?

- **Solution.** Let X be the number of clouds that show splendid growth.
- Clearly, X is a binomial random variable.
- $P(X = 35) = \binom{60}{35}(0.58)^{35}(0.42)^{25} \approx 0.1039$

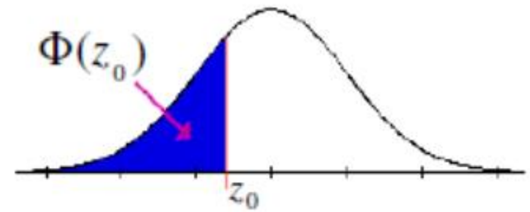
Normal Approximation

Solution: Let X be the number of clouds that show splendid growth. Then $E(X) = 60(0.58) = 34.80$ and $\sigma_X = \sqrt{60(0.58)(1-0.58)} = 3.82$. By correction for continuity and De Moivre-Laplace theorem,

$$
\begin{aligned}
P(X = 35) &\approx P(34.5 < X < 35.5) \\
&= P\Big(\frac{34.5 - 34.80}{3.82} < \frac{X - 34.80}{3.82} < \frac{35.5 - 34.80}{3.82}\Big) \\
&= P\Big(-0.08 < \frac{X - 34.8}{3.82} < 0.18\Big) \approx \frac{1}{\sqrt{2\pi}} \int_{-0.08}^{0.18} e^{-x^2/2}\,dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0.18} e^{-x^2/2}\,dx - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-0.08} e^{-x^2/2}\,dx \\
&= \Phi(0.18) - \Phi(-0.08) = 0.5714 - 0.4681 = 0.1033.
\end{aligned}
$$

Table 1 Area under the Standard Normal Distribution to the Left of z_0: Negative z_0

$$\Phi(z_0) = P(Z \le z_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_0} e^{-x^2/2}\, dx$$

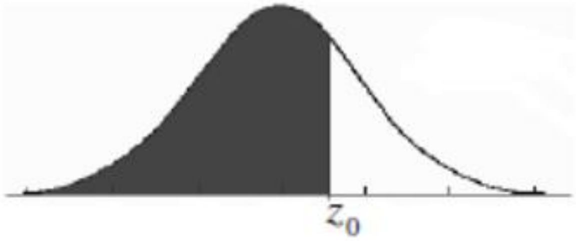


Note that for $z_0 \le -3.90$, $\Phi(z_0) = P(Z \le z_0) \approx 0$.

z_0	0	1	2	3	4	5	6	7	8	9
−3.8	.0001	0001	0001	0001	0001	0001	0001	0001	0001	0001
−3.7	.0001	0001	0001	0001	0001	0001	0001	0001	0001	0001
−3.6	.0002	0002	0001	0001	0001	0001	0001	0001	0001	0001
−3.5	0002	0002	0002	0002	0002	0002	0002	0002	0002	0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Table 2 Area under the Standard Normal Distribution to the Left of z_0: Positive z_0

$$\Phi(z_0) = P(Z \le z_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_0} e^{-x^2/2}\, dx$$



z_0	0	1	2	3	4	5	6	7	8	9
.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	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389

The Expectation and Variance of Normal Distributions

- Since the integrand $xe^{-x^2/2}$ is a finite odd function, we have

$$E(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-x^2/2}\,dx = 0.$$

- For the second moment and the variance,

$$E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,dx.$$

- By integration by parts, we have

$$\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\, dx = -xe^{-x^2/2}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2} dx$$
$$= 0 + \sqrt{2\pi} = \sqrt{2\pi}.$$

$$u = x, dv = xe^{-x^2/2}dx$$
$$du = dx, v = -e^{-x^2/2}$$

- Thus, $E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = \sqrt{2\pi}/\sqrt{2\pi} = 1$
- Moreover, $\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = 1$ and $\sigma_X = \sqrt{\mathrm{Var}(X)} = 1$
- The expected value of a standard normal random variable is 0.
- Its stand deviation is 1.

Definition of Normal Random Variables (Gaussian distribution)

- **Definition.** A random variable X is called normal, with parameters μ and σ, if its density function is given by

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

- If X is a normal random variable with parameters μ and σ, we write $X \sim N(\mu, \sigma^2)$.

Graphical Illustration

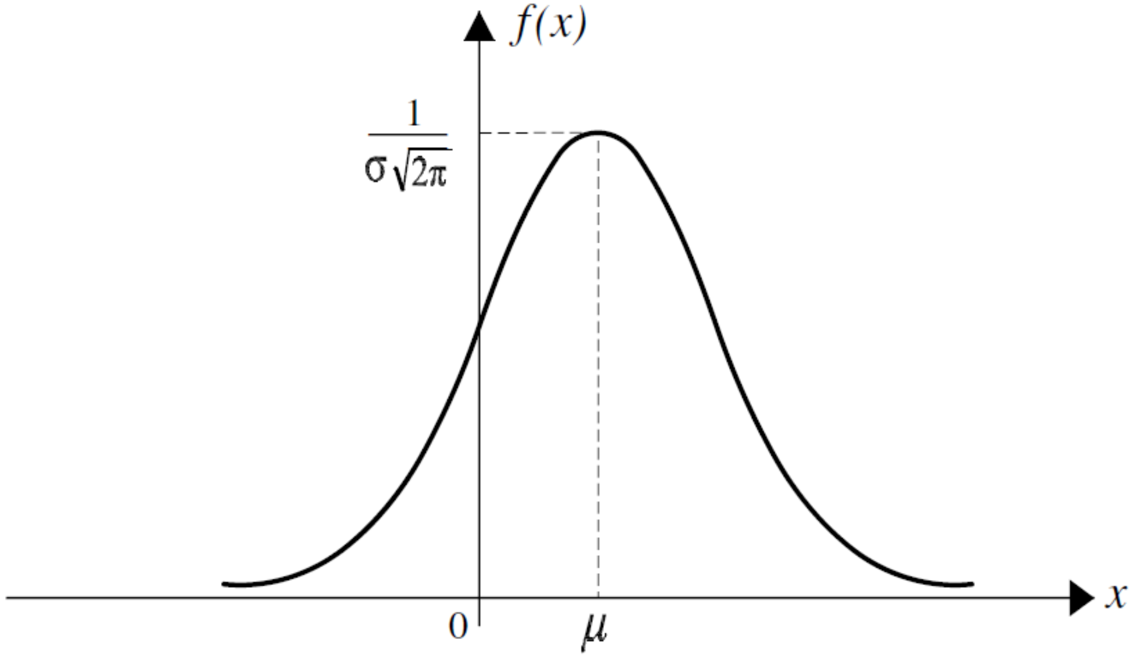
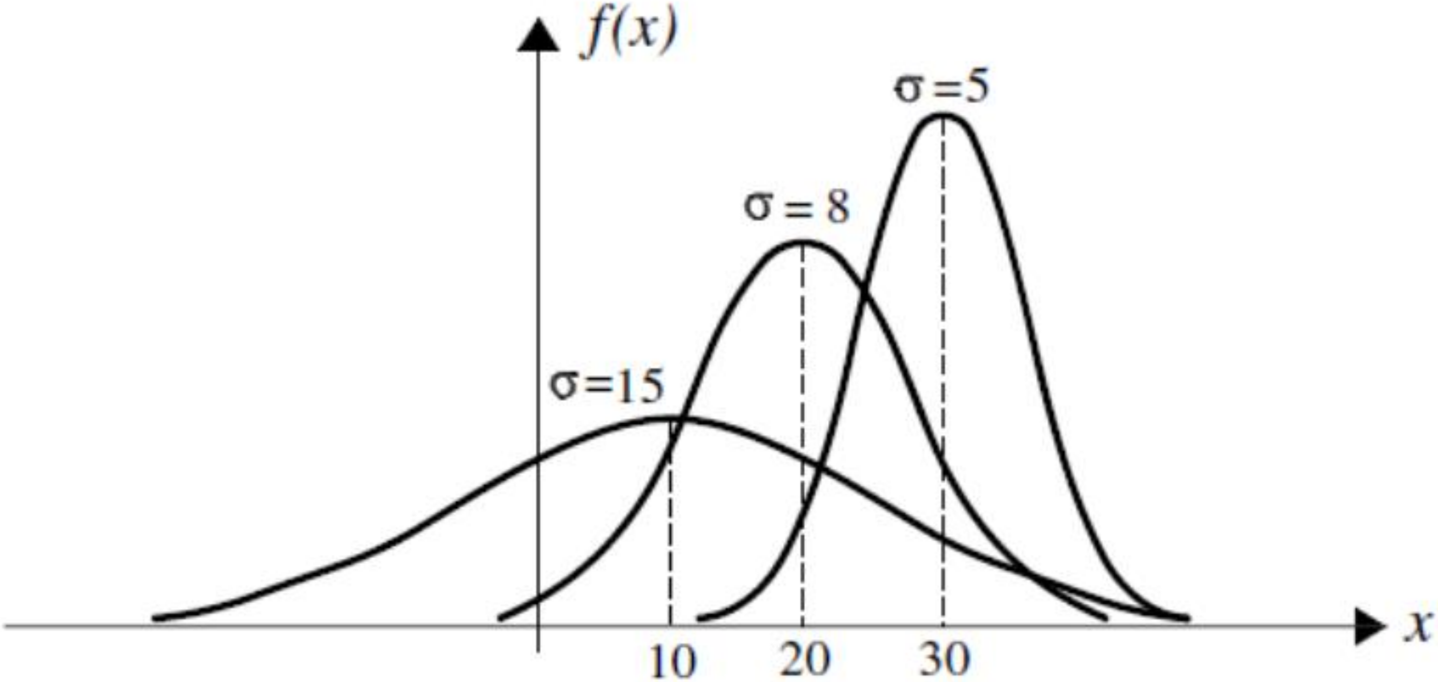


Figure 7.8 Density of $N(\mu, \sigma^2)$.

Normal Curves with Different Parameters



Standardization

- The random variable

$N(\mu, \sigma^2)$

$$X^* = \frac{X - \mu}{\sigma}$$

standard normal

$N(0, 1)$

is called the standardized X.

Lemma 7.1

- If $X \sim N(\mu, \sigma^2)$, then $Z = (X - \mu)/\sigma$ is $N(0, 1)$.
- That is, if $X \sim N(\mu, \sigma^2)$, the standardized X is $N(0, 1)$.
- **Proof.**
- We show the distribution function of Z, i.e.

$$P(Z \le x) = P\left(\frac{X - \mu}{\sigma} \le x\right) = P(X \le \sigma x + \mu)$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\sigma x + \mu} \exp\left[\frac{-(t - \mu)^2}{2\sigma^2}\right] dt.$$

- Let $y = (t - \mu)/\sigma$; then $dt = \sigma\, dy$ and we get

$$P(Z \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} \sigma\, dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$$

- The parameters μ and σ that appear in the formula of the density function are its **expected value** and **standard deviation**, respectively.

$$Z = (X - \mu)/\sigma \quad \text{is} \quad N(0,1)$$

$$X = \sigma Z + \mu$$

$$E(X) = E(\sigma Z + \mu) = \sigma E(Z) + \mu = \mu$$

$$\mathrm{Var}(X) = \mathrm{Var}(\sigma Z + \mu) = \sigma^2 \mathrm{Var}(Z) = \sigma^2$$

Example 7.6

- Suppose that a Scottish soldier's chest size is normally distributed with mean 39.8 and standard deviation 2.05 inches, respectively.
- What is the probability that of 20 randomly selected Scottish soldiers, five have a chest of at least 40 inches?

- Solution.
- Let p be the probability that a randomly selected Scottish soldier has a chest of 40 or more inches.
- If X is the normal random variable with mean 39.8 and standard deviation 2.05.

$$p = P(X \geq 40) = P\left(\frac{X - 39.8}{2.05} \geq \frac{40 - 39.8}{2.05}\right)$$

Z

$$= P\left(\frac{X - 39.8}{2.05} \geq 0.10\right)$$

$$= P(Z \geq 0.10) = 1 - \Phi(0.1) \approx 1 - 0.5398 \approx 0.46.$$

N(0,1)

- Therefore, the probability that of 20 randomly selected Scottish soldiers, five have a chest of at least 40 inches is

$$\binom{20}{5}(\underbrace{0.46}_{p})^5(0.54)^{15} \approx 0.03.$$

Example 7.9 (Investment)

- The annual rate of return for a share of a specific stock is a normal random variable with mean 10% and standard deviation 12%.
- Ms. Couture buys 100 shares of the stock at a price of $60 per share.
- What is the probability that after a year her net profit from that investment is at least $750?
  - Ignore transaction costs and assume that there is no annual dividend.

Solution of Example 7.9

- Let r be the rate of return of this stock.
- The random variable r is normal with $\mu = 0.10$ and $\sigma = 0.12$.
- Let X be the price of the total shares of the stock that Ms. Couture buys this year.
- We are given that $X = 6000$.
- Let Y be the total value of the shares next year.
- The desired probability is

$$P(Y - X \geq 750) = P\left(\frac{Y - X}{X} \geq \frac{750}{X}\right) = P\left(r \geq \frac{750}{6000}\right)$$

$$= P(r \geq 0.125) = P\left(Z \geq \frac{0.125 - 0.10}{0.12}\right)$$

$$= P(Z \geq 0.21) = 1 - P(Z < 0.21)$$

$$= 1 - \Phi(0.21) = 1 - 0.5832 = 0.4168.$$

Z is a standard normal rv.

A Few Final Words

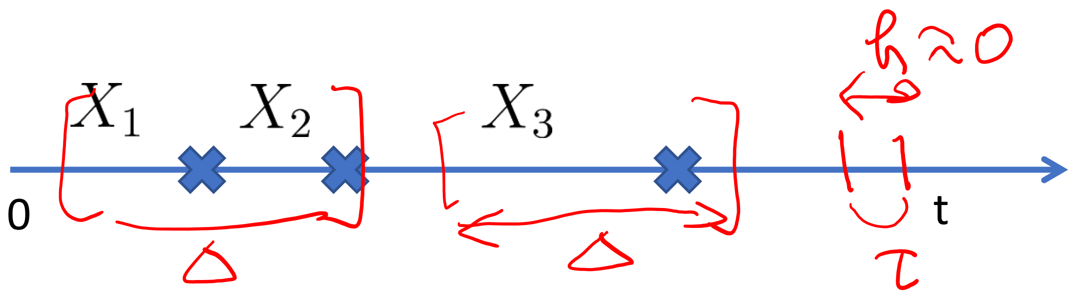
- Normal distributions appear in many applications in engineering, statistics and mathematics
- Richard von Mises, “Probability, Statistics and the Truth”, Dover Publications, 1981

The normal curve represents the distribution in all cases where a final collective is formed by combination of a very large number of initial collectives, the attribute in the final collective being the sum of the results in the initial collectives. The original collectives are not necessarily simple alternatives as they are in Bernoulli’s problem. It is not even necessary for them to have the same attributes or the same distributions. The only conditions are that a very great number of collectives are combined and that the attributes are mixed in such a way that the final attribute is the sum of all the original ones. Under these conditions the final distribution is always represented by a normal curve.

Section 7.3

Exponential Random Variables

Interarrival Times of Poisson Processes



- Recall the three assumptions of a Poisson process $\{N(t): t \geq 0\}$.
  - Stationarity
  - Independent increments
  - Orderliness

$P(N(\tau)=1) \approx \lambda \cdot \tau + \underline{o(\tau)}$

$P(N(\tau)=0) \approx 1 - \lambda\tau + o(\tau)$

$P(N(\tau) \geq 2) \approx o(\tau)$

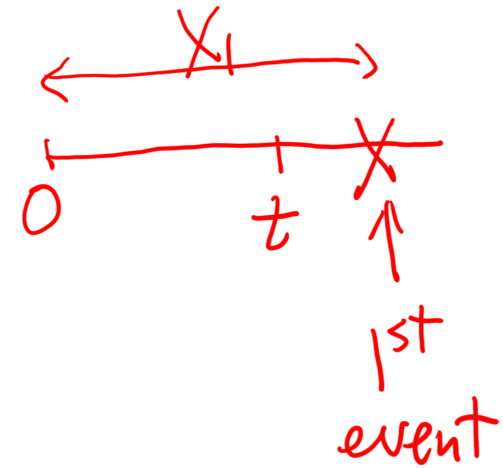
- $\{X_1, X_2, X_3, \cdots\}$ is called the sequence of interarrival times of the Poisson process $\{N(t): t \geq 0\}$.

$P(N(\Delta)=i)$

- What is the distribution of the interarrival times? $= \dfrac{e^{-\lambda|\Delta|}(\lambda|\Delta|)^i}{i!}$

Interarrival Times of Poisson Processes

- $\{X_1, X_2, X_3, \cdots\}$ is called the sequence of interarrival times of the Poisson process $\{N(t): t \geq 0\}$.



- Let $\lambda = E[N(1)]$.
- Then,

$$P(N(t) = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}.$$

- Therefore,

λ

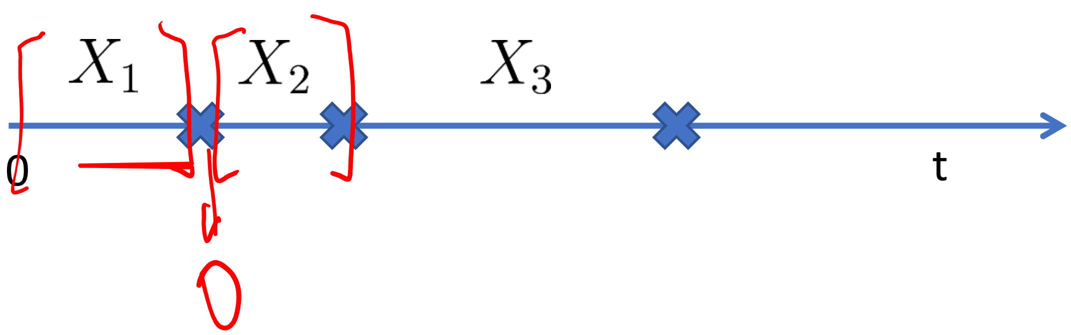
$$P(X_1 > t) = P(N(t) = 0) = e^{-\lambda t}.$$

CDF

$$P(X_1 \leq t) = 1 - P(X_1 > t) = 1 - e^{-\lambda t}.$$

λ

- Since a Poisson process is stationary and possesses independent increments, at any time t, the process probabilistically starts all over again
- Hence, the inter-arrival time of any two consecutive events has the same distribution as X_1.
- That is, $\{X_1, X_2, X_3, \cdots\}$ is a sequence of independent, and identically distributed exponential random variables



Definition of Exponential Random Variables

- Let

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$f(t) = F'(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{7.2}$$

- f is a density function since

$$\int_0^\infty f(t)\,dt = 1.$$

- **Definition.** A continuous random variable X is called exponential with parameter $\lambda > 0$ if its density function is given by (7.2).

Graphical Illustration

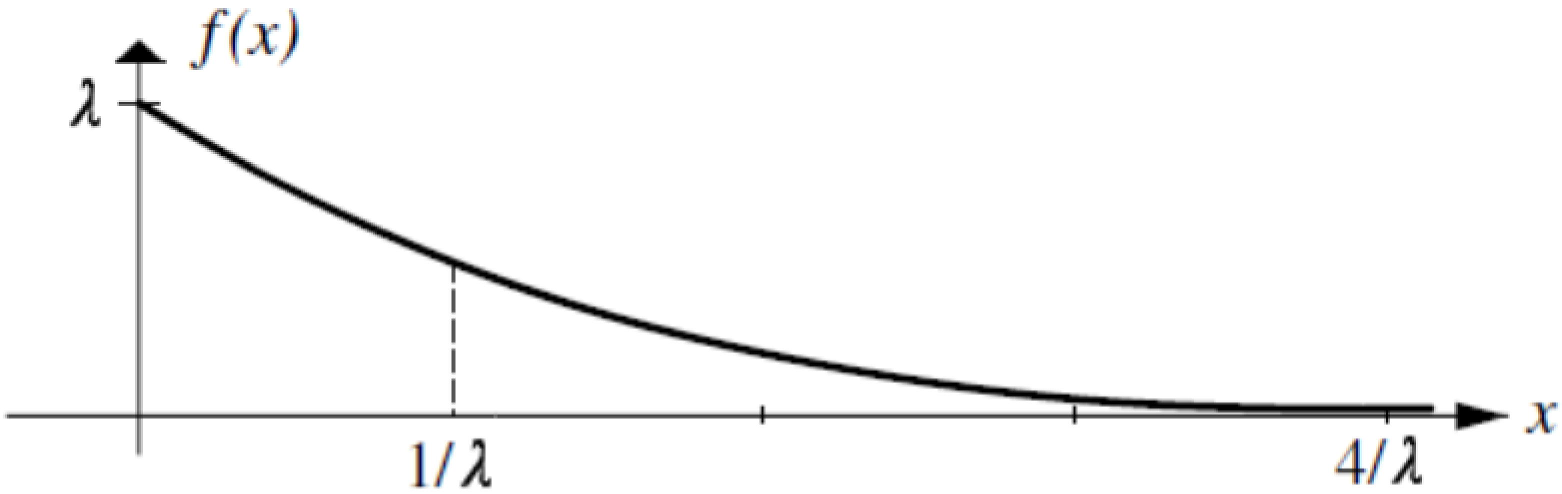


Figure 7.10 Exponential density function with parameter λ.

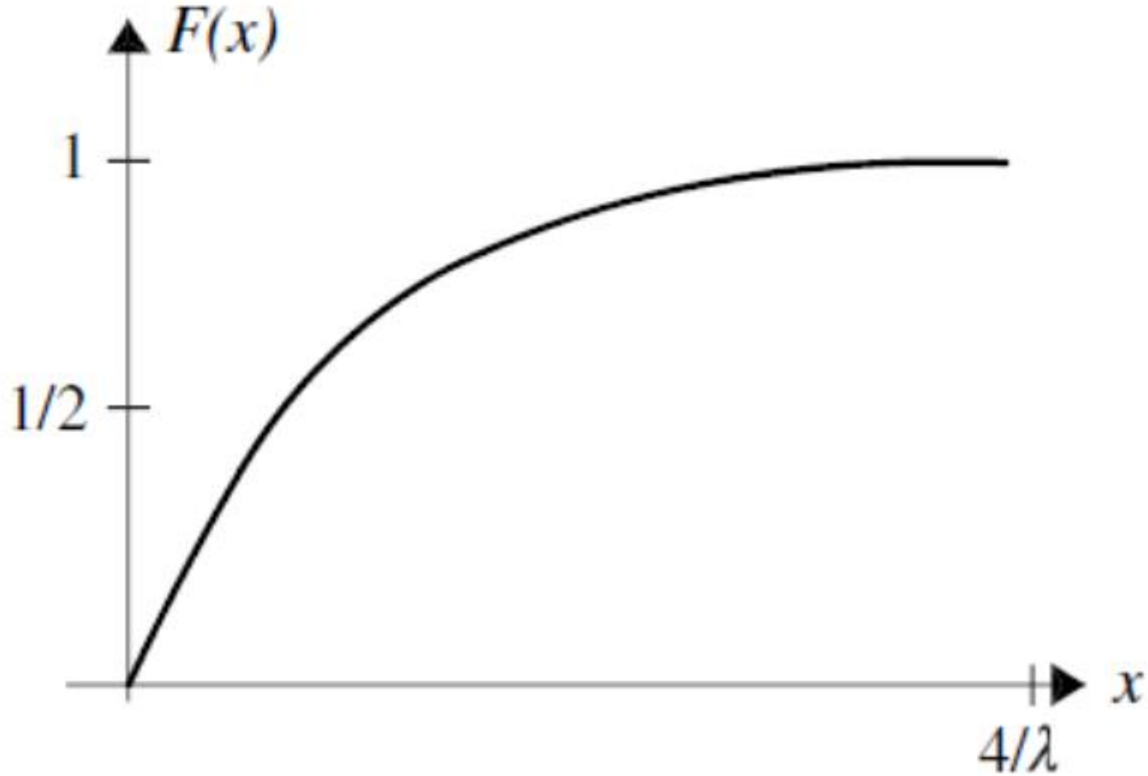


Figure 7.11 Exponential distribution function.

Expectation and Variance

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_0^{\infty} x(\lambda e^{-\lambda x})\, dx.$$

Using integration by parts with $u = x$ and $dv = \lambda e^{-\lambda x}\, dx$, we obtain

$$E(X) = \left[-xe^{-\lambda x}\right]_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\, dx = 0 - \left[\frac{1}{\lambda} e^{-\lambda x}\right]_0^{\infty} = \frac{1}{\lambda}.$$

A similar calculation shows that

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_0^{\infty} x^2(\lambda e^{-\lambda x})\, dx = \frac{2}{\lambda^2}.$$

$u = x^2$, $dv = \lambda e^{-\lambda x} dx$

$$\boxed{E(X) = \sigma_X = \frac{1}{\lambda}, \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}.}$$

3 months

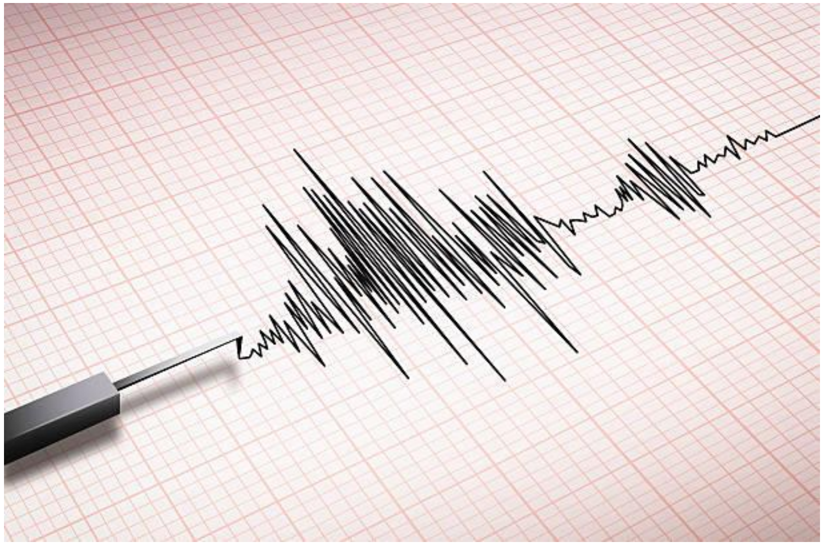
1 month

1/3 earthquakes

Example 7.10

- Suppose that every three months , on average, an earthquake occurs in California.
- What is the probability that the next earthquake occurs after three but before seven months?

$3 < X < 7$



Solution of Example 7.10

- Let X be the time (in months) until the next earthquake.
- It can be assumed that X is an exponential random variable with $1/\lambda = 3$ or $\lambda = 1/3$.
- The distribution function of X is

$$F(t) = P(X \leq t) = 1 - e^{-t/3} \quad \text{for } t > 0.$$

- Thus,

$$P(3 < X < 7) = F(7) - F(3) = (1 - e^{-7/3}) - (1 - e^{-1}) \approx 0.27.$$

Memoryless Property

- A random variable is memoryless if, for all $s, \ t$ greater than or equal to zero,

$$P(X > s + t \mid X > t) = P(X > s).$$

- Exponential is the only memoryless continuous random variable.
- Geometric is the only memoryless discrete random variable.

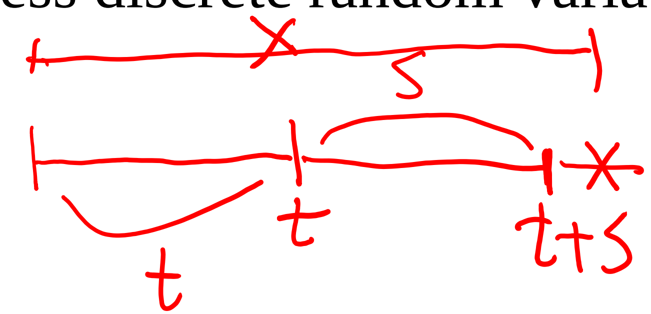


Illustration of the Memoryless Property

- $P(X > t + s \mid X > t) = P(X > s)$. (7.3)
- If, for example, X is the lifetime of some type of instrument, then (7.3) means that there is no deterioration with age of the instrument.
- The probability that a new instrument will last more than s years is the same as the probability that a used instrument that has lasted more than t years will last at least another s years.
- In other words, the probability that such an instrument will deteriorate in the next s years does not depend on the age of the instrument.

- To show that an exponential distribution is memoryless, note that (7.3) is equivalent to

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

and

$$P(X > s + t) = P(X > s)(X > t). \tag{7.4}$$

- Now since

$$P(X > s + t) = 1 - \left[1 - e^{-\lambda(s+t)}\right] = e^{-\lambda(s+t)},$$
$$P(X > s) = 1 - \left(1 - e^{-\lambda s}\right) = e^{-\lambda s},$$

and

$$P(X > t) = 1 - \left(1 - e^{-\lambda t}\right) = e^{-\lambda t},$$

(7.4) follows.

Example 7.12

- The lifetime of a TV tube (in years) is an exponential random variable with mean 10. $\Rightarrow \lambda = 1/10$
- If Jim bought his TV set 10 years ago, what is the probability that its tube will last another 10 years?
- **Solution.**
- Let X be the lifetime of the tube.
- Since X is an exponential random variable, there is no deterioration with age of the tube.
- Hence

$= 20 - 10$

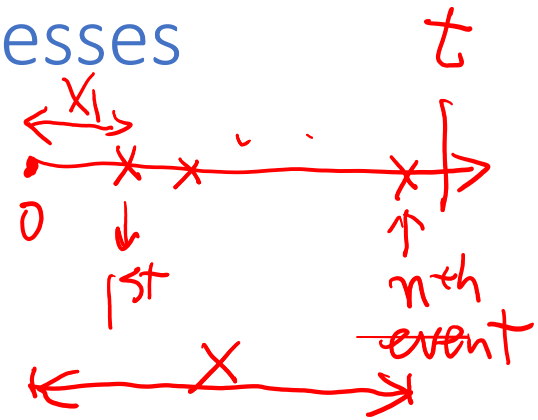
$$P(X > 20 \mid X > 10) = P(X > 10) = 1 - P(X \leq 10)$$

$$= 1 - (1 - e^{-\left(\frac{1}{10}\right)10}) \approx 0.37.$$

Section 7.4

Gamma distributions

The n-th Event Time of Poisson Processes



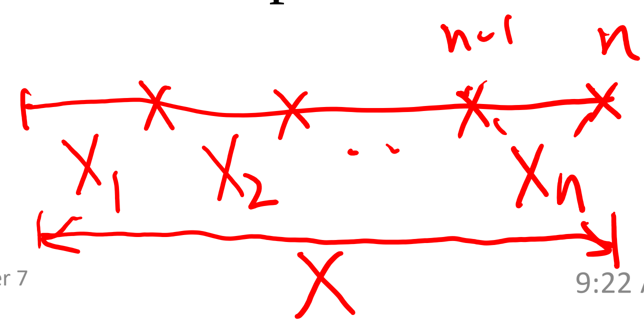
- Let $\{N(t): t \geq 0\}$ be a Poisson process.
- Let X be the time that the n-th event occurs.
- Then, X is a gamma random variable.

$$F(t) = P(X \leq t) = P(N(t) \geq n) = \sum_{i=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^i}{i!}$$

$N(0,t)$

- Recall that the interarrival time X_i between the $(i-1)$-th event and the i-th event forms an i.i.d. sequence of exponential random variables.
- Then, a gamma random variable is a sum of exponential random variable, i.e.

$$X = X_1 + X_2 + \cdots + X_n.$$



Differentiating F, the density function f is obtained:

$$
\begin{aligned}
f(t) &= \sum_{i=n}^{\infty}\left[-\lambda e^{-\lambda t}\frac{(\lambda t)^i}{i!} + \lambda e^{-\lambda t}\frac{(\lambda t)^{i-1}}{(i-1)!}\right] \\
&= \sum_{i=n}^{\infty} -\lambda e^{-\lambda t}\frac{(\lambda t)^i}{i!} + \left[\lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{i=n+1}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^{i-1}}{(i-1)!}\right] \\
&= \left[-\sum_{i=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^i}{i!}\right] + \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!} + \left[\sum_{i=n}^{\infty}\lambda e^{-\lambda t}\frac{(\lambda t)^i}{i!}\right] \\
&= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}.
\end{aligned}
$$

$\frac{d}{dx}[f(x)g(x)] = f' \cdot g + f \cdot g'$

Gamma Functions – Extension of Factorials

- Recall that factorial $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$.
- Thus, $n! = n \cdot (n-1)!$
- We want to define a function $\Gamma(r)$, such that $\Gamma(r+1) = r \cdot \Gamma(r)$, where $r > 1$ and may not be an integer.

Gamma Functions

- To extend the definition of gamma density function from parameters (n, λ) to (r, λ), where $r > 0$ may be a real number, we introduce gamma functions.

$$\Gamma(r) = \int_0^\infty t^{r-1} e^{-t}\, dt$$

- Property:

$$\Gamma(r+1) = r\Gamma(r), \qquad r > 1$$

$$\Gamma(r+1) = \int_0^\infty t^r e^{-t}\, dt = [-t^r e^{-t}]_0^\infty + r\int_0^\infty t^{r-1} e^{-t}\, dt$$

$$= r\int_0^\infty t^{r-1} e^{-t}\, dt = r\Gamma(r).$$

Integration by parts

$u = t^r$

$dv = e^{-t} dt$

- To show that $\Gamma(n+1) = n!$ for integer n, note that

$$\Gamma(1) = \int_0^\infty e^{-t}\, dt = 1$$
$$\Gamma(2) = (2-1)\Gamma(2-1) = 1 = 1!$$
$$\Gamma(3) = (3-1)\Gamma(3-1) = 2 \cdot 1 = 2!$$
$$\Gamma(4) = (4-1)\Gamma(4-1) = 3 \cdot 2 \cdot 1 = 3!$$

Definition of Gamma Distributions

Definition *A random variable X with probability density function*

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{r-1}}{\Gamma(r)} & \textit{if } x \geq 0 \\ 0 & \textit{elsewhere} \end{cases}$$

is said to have a ***gamma*** *distribution with parameters* (r, λ), $\lambda > 0$, $r > 0$.

where

$$\Gamma(r) = \int_0^\infty t^{r-1} e^{-t}\, dt.$$

Graphical Illustration

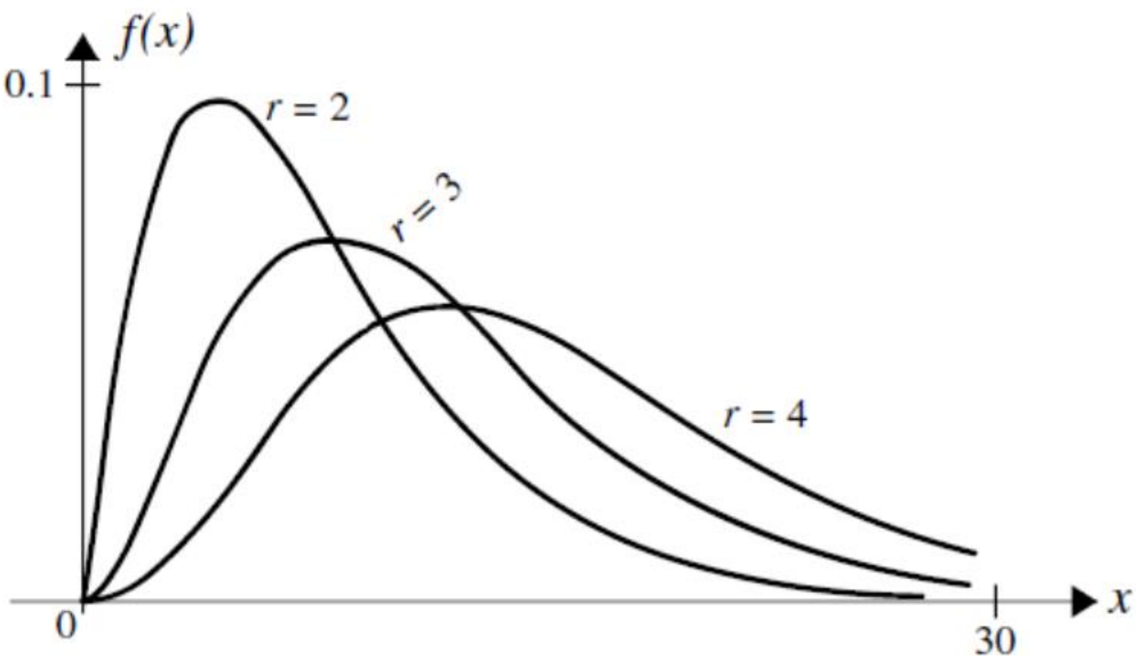


Figure 7.12 Gamma densities for $\lambda = 1/4$.

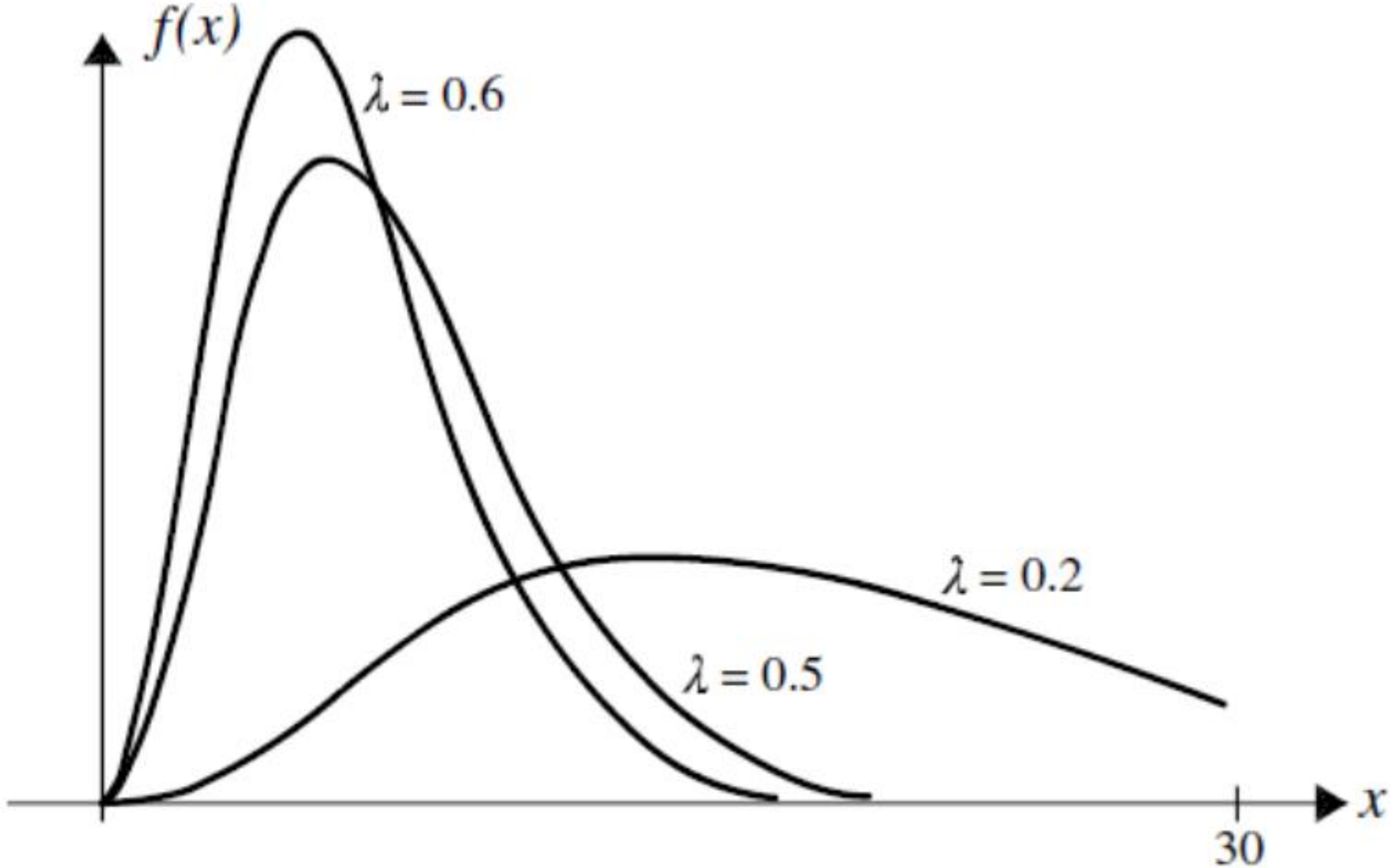
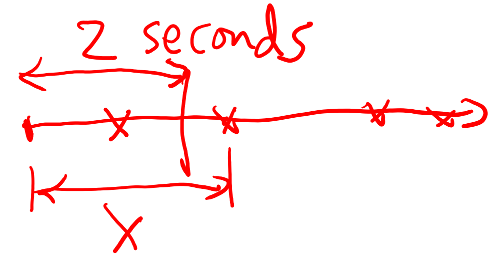


Figure 7.13 Gamma densities for $r = 4$.

Example 7.14



- Suppose that, on average, the number of β-particles emitted from a radioactive substance is four every second.
- What is the probability that it takes at least 2 seconds before the next two β-particles are emitted?
- **Solution.**
- Let $N(t)$ denote the number of β-particles emitted from a radioactive substance in $[0, t]$.
- It is reasonable to assume that $\{N(t): t \geq 0\}$ is a Poisson process.
- Let 1 second be the time unit; then $\lambda = E[N(1)] = 4$.
- X, the time between now and when the second β-particle is emitted, has a gamma distribution with parameters (2, 4).

- Therefore,

u dv

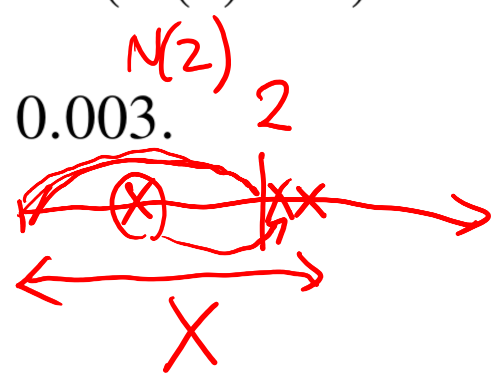
$$P(X \geq 2) = \int_2^\infty \frac{4e^{-4x}(4x)^{2-1}}{\Gamma(2)}dx = \int_2^\infty 16xe^{-4x}\,dx$$

$$= \left[-4xe^{-4x}\right]_2^\infty - \int_2^\infty (-4x^{-4x})dx = 8e^{-8} + e^{-8} \approx 0.003.$$

- An alternative solution is

$$P(X \geq 2) = P(N(2) \leq 1) = P(N(2) = 0) + P(N(2) = 1)$$

$$= \frac{e^{-8}(8)^0}{0!} + \frac{e^{-8}(8)^1}{1!} = 9e^{-8} \approx 0.003.$$



The Moments

- Let X be a gamma random variable with parameters (r, λ).

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

- Let $t = \lambda x$; then $dt = \lambda\, dx$, so

$$E(X^n) = \frac{\lambda^r}{\Gamma(r)} \int_0^\infty \frac{t^{n+r-1}}{\lambda^{n+r-1}} e^{-t} \frac{1}{\lambda} dt$$
$$= \frac{\lambda^r}{\Gamma(r)\lambda^{n+r}} \int_0^\infty t^{n+r-1} e^{-t}\, dt = \frac{\Gamma(n+r)}{\Gamma(r)\lambda^n}.$$

$\Gamma(r+1) = r\,\Gamma(r)$

For $n = 1$, this gives

$$E(X) = \frac{\Gamma(r+1)}{\Gamma(r)\lambda} = \frac{r\Gamma(r)}{\lambda\Gamma(r)} = \frac{r}{\lambda}.$$

For $n = 2$,

$$E(X^2) = \frac{\Gamma(r+2)}{\Gamma(r)\lambda^2} = \frac{(r+1)\Gamma(r+1)}{\lambda^2\Gamma(r)} = \frac{(r+1)r\Gamma(r)}{\lambda^2\Gamma(r)} = \frac{r^2+r}{\lambda^2}.$$

Thus

$$\mathrm{Var}(X) = \frac{r^2+r}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}.$$

The Expectation and Variance of Gamma Random Variables

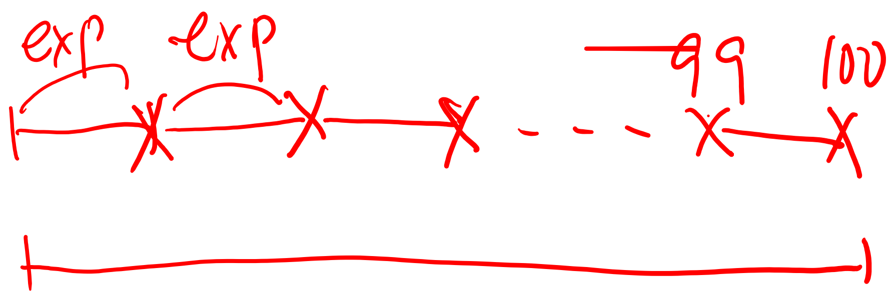
For a gamma random variable with parameters r and λ,

$$E(X) = \frac{r}{\lambda}, \qquad \operatorname{Var}(X) = \frac{r}{\lambda^2}, \qquad \sigma_X = \frac{\sqrt{r}}{\lambda}.$$

$X = X_1 + X_2 + \cdots + X_n$

Example 7.15

- There are 100 questions in a test.
- Suppose that, for all $s > 0$ and $t > 0$, the event that it takes t minutes to answer one question is independent of the event that it takes s minutes to answer another one.
- If the time that it takes to answer a question is exponential with mean 1/2, find the distribution, the average time, and the standard deviation of the time it takes to do the entire test.



Solution of Example 7.15

- Let X be the time to answer a question and $N(t)$ the number of questions answered by time t.
- Then $\{N(t): t \geq 0\}$ is a Poisson process at the rate of $\lambda = \frac{1}{E(X)} = 2$ per minute.
- Therefore, the time that it takes to complete all the questions is gamma with parameters (100, 2).
- The average time to finish the test is $\frac{r}{\lambda} = \frac{100}{2} = 50$ minutes.
- The standard deviation is $\sqrt{r/\lambda^2} = \sqrt{100/4} = 5$.

Homework 9

- Section 7.2: 23, 24, 29, 30
- Section 7.3: 3, 6, 16
- Section 7.4: 6, 9

- Due date: 5 pm, Wednesday, May 17, 2023.

- Sections 7.5 and 7.6 are skipped.