# SECTION 10: SPECIAL PROBABILITY DENSITIES

In this section the most prominent probability densities and some characteristics of them shall be given. These densities /distributions are widely used in applications.

## 10.1. The Uniform Distribution

If any random variable takes value in an interval with same chance, the random variable has uniform distribution in the interval (a, b). it is continuous version of the discrete uniform distribution.

**Definition 10.1.1** A random variable has uniform distribution and it is referred to as a continuous random variable if only if its probability density is given by:

$$f(x) = \frac{1}{b-a}, \quad a < x < b$$
$$= 0, \quad otherwise$$

**Theorem 10.1.1** The mean and variance of the uniform distribution are given by:

$$E(X) = \frac{a+b}{2}, \qquad V(X) = \frac{(b-a)^2}{12}$$

**Example 1:** In certain experiments, the error made in determining the density of a substance is a random variable having a uniform density with a=-0.015 and b=0.015. Find the probabilities that such an error will

a) be between -0.002 and 0.003
b) exceed 0.005 in absolute value.

**Solution:** If *X* has uniform distribution (-0.015, 0.015), the probability density function of X is:

$$f(x) = \frac{1}{0.03}, \quad -0.015 < x < 0.015$$
$$= 0, \quad otherwise$$

**a)** $P(-0.002 < X < 0.003) = \int_{-0.002}^{0.003} \frac{1}{0.03} dx = \frac{x}{0.03}\Big|_{-0.002}^{0.003} = \frac{0.003 + 0.002}{0.03} = \frac{1}{6}$

**b)** $P(|X| \geq 0.005) = 1 - P(|X| < 0.005) = 1 - \int_{-0.005}^{0.005} \frac{1}{0.03} dx = \frac{x}{0.03}\Big|_{-0.005}^{0.005} = \frac{2 \times 0.005}{0.03} = \frac{1}{3}$

## 10.2. The Exponential Distribution

The distribution is used for modeling of the random variable which is

- the lifetime of an individual picked at random from some biological population,
- the time decay of a radioactive atom,
- the CPU time for a specific program or processing in the computer,
- the waiting time for a customer in serving line etc.

**Definition 10.2.1** A random variable has exponential distribution and it is referred to as a continuous random variable if only if its probability density is given by:

$$
\begin{aligned}
f(x) &= \frac{1}{\theta} e^{-x/\theta}, \quad x > 0 \\
&= 0, \qquad\qquad x \le 0
\end{aligned}
$$

**Theorem 10.2.1** The mean and variance of the exponential distribution are given by:

$$E(X) = \theta, \qquad V(X) = \theta^2$$

**Example 2:** Suppose a certain solid-state component has a lifetime or failure time (in hours) X~EXP(θ=100). What is the probability that the component will last at least 50 hours?

**Solution:**

$$P(X \ge 50) = 1 - P(X < 50) = 1 - \int_0^{50} \frac{1}{100} e^{-x/100} dx = 1 - (-e^{-0.5} + 1) = e^{-0.5} = 0.6065$$

**Theorem 10.2.2** The generation function of the exponential distribution is given by:

$$M_X(t) = \frac{1}{1 - \theta t}.$$

**Theorem 10.2.3** *X* has exponential distribution with parameter θ and for s≥0, t≥0, if only if *X* has memoryless property

$$P\left(X > t + s \middle| X > t\right) = P(X > s).$$

Proof:

$$
\begin{aligned}
P\left(X > t + s \middle| X > t\right) &= \frac{P\left(X > t + s\right)}{P\left(X > t\right)} = \frac{1 - P\left(X \le t + s\right)}{1 - P\left(X \le t\right)} = \frac{1 - F_X\left(t + s\right)}{1 - F_X\left(t\right)} \\
&= \frac{e^{-(t+s)/\theta}}{e^{-t/\theta}} = e^{-s/\theta} = P\left(X > s\right).
\end{aligned}
$$

Where $P(X > x) = 1 - F_X(x)$ is referred to as a survival function of *X*.

**Example 3:** Suppose the average lifetime of particular kind of transistor is 100 working hours, and that the life distribution is approximately exponential with the parameter θ=100.

**a)** Estimate the probability that the transistor will work for at least 50 hours.
**b)** If it is known that the transistor has worked for 50 hours, what is the probability that the transistor will work for more 50 hours.

**Solution:**
The density probability function of *X* is as follows:

$$f(x)=\frac{1}{100}e^{-x/100}, \quad x>0$$
$$=0, \qquad otherwise$$

a) $P(X>50)=\int_{50}^{\infty}\frac{1}{100}e^{-x/100}dx=-e^{-x/100}\Big|_{50}^{\infty}=e^{-50/100}=0.606$

b) Form the memoryless property of the distribution, the probability is:

$$P\left(X>50+50\middle|X>50\right)=\frac{P\left(X>100\right)}{P\left(X>50\right)}=\frac{e^{-100/100}}{e^{-50/100}}=e^{-50/100}=P\left(X>50\right)=0.606\,.$$

### 10.3. The Normal Distribution

The normal distribution is known as a cornerstone of modern statistical theory. It was investigated first in the eighteenth century when scientists observed an astonishing degree of regularity in errors of measurement. They found that the patterns (distributions) that they observed could be closely approximated by continuous curves, which they referred to as "normal curves of errors" and attributed to the laws of chance. The mathematical properties of such normal curves were first studied by Abraham de Moivre (1667-1745), Pierre Laplace (1749-1745), and Karl Gauss (1777-1855) (Miller and Miller, 2004).

**Definition 10.3.1** A random variable has normal distribution and it is referred to as a continuous random variable if only if its probability density is given by:

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty<x<\infty$$
$$=0, \qquad otherwise$$

**Theorem 10.3.1** The mean and variance of the normal distribution are given by:

$$E(X)=\mu, \qquad V(X)=\sigma^2$$

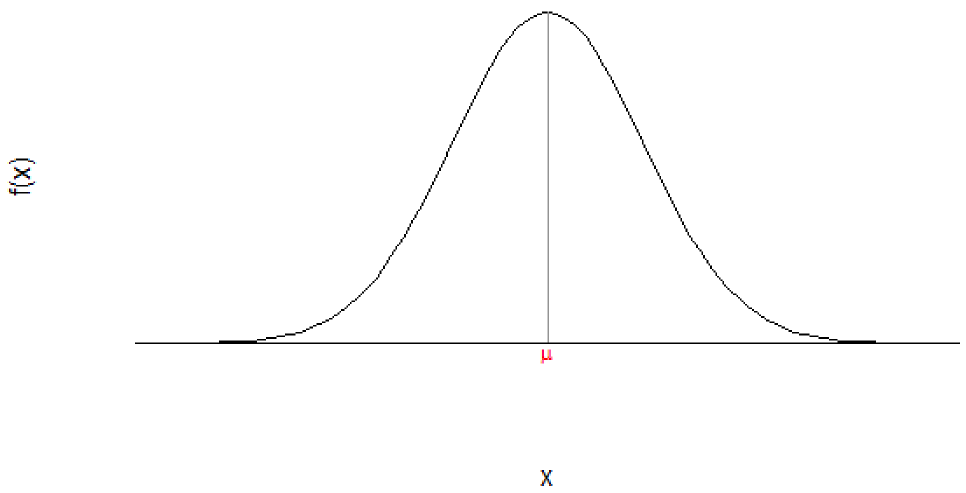


**Figure 10.3.1** Graph of normal distribution.

**Theorem 10.3.2** The generation function of the normal distribution is given by:

$$M_X(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}.$$

Let $X$ be a random variable which has normal distribution with parameters μ and $\sigma^2$, the $Z = \dfrac{X-\mu}{\sigma}$ standard variable has normal distribution with parameters $\mu$=0 and $\sigma^2$=1 and the distribution is referred to as standard normal distribution.

**Definition 10.3.2** A random variable has standard normal distribution and it is referred to as a continuous random variable if only if its probability density is given by:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty$$
$$= 0, \qquad otherwise$$

Where the mean and variance of $Z$ is 0 and 1 respectively.

**Example 4:** Find the probabilities that a random variable having the standard normal distribution will take on a value

- **a)** less than 1.72.
- **b)** less than -0.88.
- **c)** between 1.30 and 1.75.
- **d)** between -0.25 and 0.45.

**Solution:** Using standard normal distribution table,

a) $$\begin{aligned} P(Z<1.72) &= 0.5+P(0<Z<1.72) \\ &= 0.5+0.4573=0.9573 \end{aligned}.$$

b) $$\begin{aligned} P(Z<-0.88) &= 0.5-P(0<Z<0.88) \\ &= 0.5-0.3106=0.1894 \end{aligned}.$$

c) $$\begin{aligned} P(1.30<Z<1.75) &= P(0<Z<1.75)-P(0<Z<1.30) \\ &= 0.4599-0.4032=0.0567 \end{aligned}.$$

d) $$\begin{aligned} P(-0.25<Z<0.45) &= P(0<Z<0.25)+P(0<Z<0.45) \\ &= 0.0987+0.1736=0.2723 \end{aligned}$$

**Example 5:** Suppose that the amount of cosmic radiation to which a person is exposed when flying by jet across the United States is a random variable having normal distribution with mean of 4.35 mrem and a standard deviation of 0.59 mrem. What is the probability that a person will exposed to more than 5.20 mrem of cosmic radiation on such a flight?

**Solution:**

$$\begin{aligned} P(X>5.20) &= P\left(\frac{X-\mu}{\sigma}>\frac{5.20-\mu}{\sigma}\right)=P\left(Z>\frac{5.20-4.35}{0.59}\right) \\ &= P\left(Z>1.44\right)=0.5-P\left(0<Z<1.44\right) \\ &= 0.5-0.4251=0.0749 \end{aligned}$$

### 10.4. The Normal Approximation to the Binomial Distribution

First of all, we will give a short information about **Central Limit Theorem:**

> ***Central Limit Theorem:*** For large sample size, the mean $\bar{X}$ of the a sample from a population with mean μ and standard deviation σ possesses a sampling distribution that is approximately normal-regardless of the probability distribution of the sampled population. For the large the sample size, the approximation will be better.

From the theorem, the distibution of $\bar{X}$ , $\quad$ *for* $n\to\infty$ $(n\geq 30)$, $\quad \bar{X}\sim \mathrm{N}\left(\mu,\dfrac{\sigma^2}{n}\right)$, where the term of ∞ is corresponded to $n\geq 30$. Using the result of theorem, the standardized $\bar{X}$ has a standard normal distribution: $\quad \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim \mathrm{N}\left(0,1\right)$.

**Theorem 10.3.3** If X is a random variable having binomial distribution with the parameters n, p, then the distribution of

$$n \to \infty \quad Z = \frac{X - np}{\sqrt{npq}} \sim N(0,1)$$

approaches that of the standard normal distribution when n→∞.

This approximation as a result of **Central Limit Theorem** means that the probabilities computed from standard normal distribution are very close to the probabilities computed from binomial distribution when the number of trials are very large (especially *np* and *npq* are both greater than 5).

**Example 6**: Suppose that 3% of items produced by assembly line are defective. An inspector selects 100 items at random from assembly line.

**a)** Approximate the probability that exactly five defectives are selected.
**b)** Approximate the probability that less than or equally to five defectives are selected.

**Solution:**

$$E(X) = np = 100 \times 0.03 = 3$$
$$V(X) = npq = 100 \times 0.03 \times 0.97 = 2.91 \quad \sqrt{V(X)} = 1.7058$$

a) $P(X = 5) \Rightarrow P(4.5 < X < 5.5)$ is known as continuity correction.

$$P(X = 5) \Rightarrow P(4.5 < X < 5.5) = P\left(\frac{4.5-3}{1.7018} < \frac{X-3}{1.7018} < \frac{5.5-3}{1.7018}\right)$$
$$= P(0.88 < Z < 1.47) = P(0 < Z < 1.47) - P(0 < Z < 0.88)$$
$$= 0.4292 - 0.3106 = 0.1186$$

$$P(X \le 5) \Rightarrow P(X \le 5.5) = P\left(\frac{X-3}{1.7018} \le \frac{5.5-3}{1.7018}\right)$$

b)
$$= P(Z \le 1.47) = 0.5 + P(0 < Z < 1.47)$$
$$= 0.5 + 0.4292 = 0.9292$$

## EXERCISES

**Exercise 1:** The average amount of weight gained by a person over the winter months is uniformly distributed from 0 to 30lbs. Find the probability a person will gain between 10 and 15lbs during the winter months.

**Solution:**

$$f(x) = \frac{1}{30}, \quad 0 < x < 30$$
$$= 0, \qquad otherwise$$
$$P(10 < X < 15) = \int_{10}^{15} \frac{1}{30} dx = \frac{1}{30} x \Big|_{10}^{15} = \frac{1}{30}(15-10) = \frac{1}{6}$$

**Exercise 2:** The mileage (in thousands in miles) that car owners get with a certain kind of radial tire is random variable having an exponential distribution with θ=40. Find the probabilities that one of these tires will last

**a)** at least 20,000 miles,
**b)** at most 30,000 miles.

**Solution:**

$$f(x) = \frac{1}{40} e^{-x/40}, \quad x > 0$$
$$= 0, \qquad x \le 0$$

a) $P(X > 20) = \int_{20}^{\infty} \frac{1}{40} e^{-x/40}\, dx = -e^{-x/40}\Big|_{20}^{\infty} = e^{-20/40} = 0.606$

b) $P(X < 30) = \int_{0}^{30} \frac{1}{40} e^{-x/40}\, dx = -e^{-x/40}\Big|_{0}^{30} = 1 - e^{-30/40} = 0.528$

**Exercise 3:** The amount of time that a watch will run without having to be reset is a random variable having an exponential distribution with θ=120 days. Find the probabilities that such a watch will

**a)** have to be reset in less than 24 days,
**b)** not have to be reset in at least 180 days.
**c)** Find the expected value of X.

**Solution:**

$$f(x) = \frac{1}{120} e^{-x/120}, \quad x > 0$$
$$= 0, \qquad x \le 0$$

a) $P(X < 24) = \int_{0}^{24} \frac{1}{120} e^{-x/120}\, dx = -e^{-x/120}\Big|_{0}^{24} = 1 - e^{-24/120} = 0.181$

b) $P(X>180)=\int_{180}^{\infty}\frac{1}{120}e^{-x/120}\,dx=-e^{-x/120}\Big|_{180}^{\infty}=e^{-180/120}=e^{-1.5}=0.2231$

c) $\int udv=uv-\int vdu$ then

$$E(X)=\int_0^{\infty}\underset{u}{x}\underbrace{\frac{1}{120}e^{-x/120}\,dx}_{dv}=\underbrace{-xe^{-x/120}\Big|_0^{\infty}}_{0}+\int_0^{\infty}e^{-x/120}\,dx=-120e^{-x/120}\Big|_0^{\infty}=120$$

**Exercise 4:** Suppose a long series of repeated measurements of the weight of a standard kilogram yield results that normally distributed with a mean of one kilogram and standard deviation of 20 micrograms. (1 x $10^{-6}$)= 1 mcg

**a)** About what proportion of measurements are correct to within 10 micrograms?
**b)** In 100 measurements, what is the probability that more than 45 measurements will be correct to within 10 micrograms?
**c)** $Y=observed\ weight-1kg$, find the approximate average absolute size of errors, $E\left(|Y|\right)$=?

**Solution:** $X\sim N(1kg,20^2)$ the random variable X represents the observed weight of a standard kilogram yield.

**a)** $$P(-10\le X-\mu\le 10)=P\left(\frac{-10}{20}\le\frac{X-\mu}{\sigma}\le\frac{10}{20}\right)$$
$$=2P\left(0<Z\le 0.5\right)=2\times 0.1915=0.383$$

**b)** $$\sum_{y=45}^{100}\binom{100}{y}0.383^y(\underbrace{1-0.383}_{0.617})^{100-y}$$

**c)**
$$E\left(|Y|\right)=\int_{-\infty}^{+\infty}|x-1|f(x)dx=\int_{-\infty}^{1}(1-x)f(x)dx+\int_{1}^{+\infty}(x-1)f(x)dx$$
$$=\int_{-\infty}^{1}(1-x)\frac{e^{-\frac{1}{2\sigma^2}(x-1)^2}}{\sqrt{2\pi}\sigma}dx+\int_{1}^{+\infty}(x-1)\frac{e^{-\frac{1}{2\sigma^2}(x-1)^2}}{\sqrt{2\pi}\sigma}dx$$
$$=\int_{-\infty}^{0}(-u)\frac{e^{-\frac{1}{2\sigma^2}u^2}}{\sqrt{2\pi}\sigma}du+\int_{0}^{+\infty}u\frac{e^{-\frac{1}{2\sigma^2}u^2}}{\sqrt{2\pi}\sigma}du$$
$$=\int_{0}^{+\infty}u\frac{e^{-\frac{1}{2\sigma^2}u^2}}{\sqrt{2\pi}\sigma}du+\int_{0}^{+\infty}u\frac{e^{-\frac{1}{2\sigma^2}u^2}}{\sqrt{2\pi}\sigma}du$$
$$=2\int_{0}^{+\infty}u\frac{e^{-\frac{1}{2\sigma^2}u^2}}{\sqrt{2\pi}\sigma}du=\frac{2}{\sqrt{2\pi}\sigma}\left\{-\sigma^2e^{-\frac{1}{2\sigma^2}u^2}\Big|_0^{+\infty}\right\}$$
$$=\frac{2\sigma}{\sqrt{2\pi}}=\frac{2(20)}{\sqrt{2\pi}}=\frac{40}{\sqrt{2\pi}}$$

**Exercise 5:** Assume that on each day each person in a certain community with population 10,000 requires a hospital bed with probability 1/2000. At least how many beds should the hospital have so that it can accommodate all people requiring a bed on a given day with a probability of at least 0.95?

**Solution:** Using normal approximation to binomial distribution,

$$\mu = np \Rightarrow \mu = 10000 \times \tfrac{1}{2000} = 5$$

$$\sigma^2 = npq \Rightarrow \sigma^2 = 10000 \times \tfrac{1}{2000} \times \tfrac{1999}{2000} = 4.99$$

$X \sim N(5,\ 4.99)$ : X shows the number of beds required in a day for the hospital.

$$P(X \geq a) \geq 0.95 \Rightarrow P\left(\frac{X-5}{\sqrt{4.99}} \geq \frac{a-5}{\sqrt{4.99}}\right) = 0.95$$

$$P\left(0 < Z < \frac{5-a}{\sqrt{4.99}}\right) = 0.45 \Rightarrow \frac{5-a}{\sqrt{4.99}} = 1.65 \Rightarrow a = 5 - \underbrace{1.65 \times \sqrt{4.99}}_{3.68} = 1.32\,beds$$

**Exercise 6:** Suppose that the actual amount of instant coffee that a filling machine puts into "6-ounce" jars is a random variable having normal distribution with σ=0.05 ounce. If only 3 percent of the jars are to contain less than 6 ounces of coffee, what must be the fill of these jars?

**Solution:** $X \sim N(6,\ 0.05^2)$ the random variable X represents amount of instant coffee filled up 6 ounce jar.

$$P(X < a) = 0.03 \Rightarrow P\left(\frac{X-6}{0.05} < \frac{a-6}{0.05}\right) = 0.03$$

$$P\left(Z < \frac{a-6}{0.05}\right) = 0.03 \Rightarrow \frac{a-6}{0.05} = -1.88 \Rightarrow a = 6 - \underbrace{1.88 \times 0.05}_{0.094} = 5.906$$

**Exercise 7:** Suppose that the final grade of a student who takes a statistics course in a particular college is a normal distributed random variable with a mean of 70 and standard deviation of 10.

- **a)** What is the probability that a student will get at least 90 points for his final grade?
- **b)** What is the probability that a student will get less than 50 points for his final grade?
- **c)** What is the highest grade in 20% of the students have got less than F3?

**Solution:** $X \sim N(70,\ 10^2)$ the random variable X represents final grade of a student.

a) $$P(X \geq 90) = P\left(\frac{X-70}{10} \geq \frac{90-70}{10}\right) = P(Z \geq 2) = 1 - P(Z < 2)$$
$$= 0.5 - P(0 < Z < 2) = 0.5 - 0.4772 = 0.0228$$

b) $$P(X < 50) = P\left(\frac{X-70}{10} < \frac{50-70}{10}\right) = P(Z < -2) = 0.5 - P(-2 < Z < 0)$$
$$= 0.5 - P(0 < Z < 2) = 0.5 - 0.4772 = 0.0228$$

$$P(X < a) = 0.20$$

c) $$P\left(\frac{X-70}{10} < \frac{a-70}{10}\right) = 0.20$$

$$P\left(Z < \frac{a-70}{10}\right) = 0.20 \Rightarrow \frac{a-70}{10} = -0.84 \Rightarrow a = 61.6$$

**Exercise 8:** A group of students with normally distributed salaries earn an average of \$6,800 with a standard deviation of \$2,500. What proportion of students earn between \$6,500 and \$7,300?

**Solution:** $X \sim N(6,800,\ 2,500^2)$ the random variable X represents salary of a student. $P(6500 < X < 7300) = ?$

$$P(6500 < X < 7300) = P\left(\frac{6500-6800}{2500} < \frac{X-\mu}{\sigma} < \frac{7300-6800}{2500}\right)$$
$$= P\left(-0.12 < Z < 0.2\right) = P(Z < 0.2) - P(Z < -0.12)$$
$$= P(0 < Z < 0.2) + P(0 < Z < 0.12)$$
$$= 0.0793 + 0.0478 = 0.1271$$

**Exercise 9**: The diameters of steel disks produced in a plant are normally distributed with a mean of 2.5 cm and standard deviation of .02 cm. The probability that a disk picked at random has a diameter greater than 2.54 cm is about:

**Solution:** $X \sim N(2.5,\ 0.02^2)$ the random variable X represents measurement of the diameters of steel disks in a plant. $P(X > 2.54) = ?$

$$P(X > 2.54) = P\left(\frac{X-\mu}{\sigma} > \frac{2.54-2.5}{0.02}\right) = P\left(Z > \frac{0.04}{0.02}\right) = P(Z > 2) = 0.5 - P(0 < Z < 2)$$
$$= 0.5 - 0.4772 = 0.0228$$