3. Some special random variables

3.1. **Bernoulli and binomial random variables.** Suppose that in an experiment, the outcome is classified as either success or a failure. Define a random variable X as

$$X = \begin{cases} 1 \text{ if the outcome is success} \\ 0 \text{ if the outcome is failure} . \end{cases}$$

Then the probability mass function of X is given by

$$\begin{aligned} p(0) &= \mathbb{P}(X = 0) = 1 - p, \\ p(1) &= \mathbb{P}(X = 1) = p \,, \end{aligned} \tag{3.1}$$

where $p, 0 < p < 1$ is the probability that the trial is a success.

Definition 3.1. A random variable X is said to be a Bernoulli random variable if its probability mass function is given by (3.1) for some $p \in (0, 1)$.

Definition 3.2 (Binomial distribution)**.** A random variable X is said to be a binomial distribution is parameters n and p, denoted as $X \sim \mathcal{B}(n, p)$ if its probability mass function is given by

$$p(i) = \begin{cases} \binom{n}{i} p^i (1-p)^{n-1} & \text{if } i = 0, 1, \ldots, n \\ 0, \quad \text{otherwise} \end{cases} \tag{3.2}$$

Binomial random variable basically denotes the number of success achieved after carrying out n independent repetitions of the experiment. The distribution function is given by

$$\mathbb{P}(X \le i) = \sum_{k=0}^{i} \binom{n}{k} p^k (1-p)^{n-k} \quad i = 0, 1, \ldots, n \,.$$

Example 3.1. *A fair coin is tossed eight times. What is the probability of obtaining*

a) less that 4 heads, b) more than five heads.

Solution: *Let H denotes the number of heads. Then $H \sim \mathcal{B}(8, 0.5)$. Now*

$$\mathbb{P}(H \le 3) = \sum_{i=0}^{3} \mathbb{P}(H = i) = \frac{1}{2^8}\Big(1 + \binom{8}{1} + \binom{8}{2} + \binom{8}{3}\Big) = \frac{93}{256} \approx 0.363 \,,$$

$$\mathbb{P}(H > 5) = \sum_{i=6}^{8} \mathbb{P}(H = i) = \frac{1}{2^8}\Big(1 + \binom{8}{6} + \binom{8}{7}\Big) = \frac{37}{256} \approx 0.1445 \,.$$

Properties of binomial random variable: We calculate the expectation, variance of $X \sim \mathcal{B}(n, p)$. Observe that

$$\begin{aligned} \mathbb{E}[X^k] &= \sum_{i=0}^{n} i^k \binom{n}{i} p^i (1-p)^{n-i} = p \sum_{i=1}^{n} i^{k-1} i \binom{n}{i} p^{i-1} (1-p)^{n-i} \\ &= np \sum_{i=1}^{n} i^{k-1} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} = np \sum_{j=0}^{n-1} j + 1^{k-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \\ &= np \mathbb{E}[(Y+1)^{k-1}] \,, \end{aligned}$$

where $Y \sim \mathcal{B}(n-1, p)$. Hence for $k = 1$, we have $\mathbb{E}[X] = np$. For $k = 2$,

$$\mathbb{E}[X^2] = np\mathbb{E}[Y + 1] = np\{(n-1)p + 1\} \,.$$

Therefore, the variance of X is given by

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2 = np(1-p) \,.$$

Remark 3.1. If $X \sim \mathcal{B}(n,p)$, then for any $0 \le k \le n$,

$$\mathbb{P}(X=k) = \frac{(n-k+1)p}{k(1-p)}\mathbb{P}(X=k-1) \implies \mathbb{P}(X=k) \ge \mathbb{P}(X=k-1) \Leftrightarrow (n-k+1)p \ge k(1-p)$$

Hence, $\mathbb{P}(X=k)$ first increases monotonically and then decreases monotonically, reaching its largest value when k is the largest integer less than or equal to $(n+1)p$.

Example 3.2. *A baised die is thrown 30 times and the number of sixex seen is eight. If the die is thrown a further 12 times, find*

a) the probability that a six will occur exactly twice, b) the expected number of sixes, and c) the variance of the number of sixes.

Solution: *Let X be defined by " the number of sixes seen in 12 throws". Then $X \sim \mathcal{B}(12, \frac{8}{30})$. The probability that a six will occur exactly twice equals to $\mathbb{P}(X=2)$. Therefore,*

$$a)\ \mathbb{P}(X=2) = \binom{12}{2}\left(\frac{4}{15}\right)^2\left(\frac{11}{15}\right)^{10} \approx 0.211\,,$$

$$b)\ \mathbb{E}[X] = np = 12 \times \frac{4}{15} = 3.2\,, \quad c)\ \mathrm{Var}(X) = np(1-p) = 12 \times \frac{4}{15} \times \frac{11}{15} \approx 2.347\,.$$

Example 3.3. *For a binomially distributed random variable X with mean 6 and variance 4.2, find $\mathbb{P}(X \le 6)$.*

Solution: *Let $X \sim \mathcal{B}(n,p)$. From the given conditions, we have $6 = np$ and $4.2 = np(1-p)$. Solving these, we have $n = 20$ and $p = 0.3$. Hence $X \sim \mathcal{B}(20, 0.3)$. Therefore,*

$$\mathbb{P}(X \le 6) = \sum_{i=0}^{6} \mathbb{P}(X=i) = \sum_{i=0}^{6} \binom{20}{i}(0.3)^i(0.7)^{20-i} \approx 0.6080\,.$$

3.2. **Poisson random variable.** A non-negative integer valued random variable X is said to be a Poisson random variable with parameter $\lambda > 0$, denoted as $X \sim \mathcal{P}(\lambda)$, if probability mass function is given by

$$p(i) = \mathbb{P}(X=i) = e^{-\lambda}\frac{\lambda^i}{i!} \quad i = 0, 1, \dots$$

The Poisson random variable has a wide range of applications in diverse areas, For example, the number of misprints on a page of a book, the number of phone calls received by a call center follows a Poisson distribution. The Poisson distribution is actually a limiting case of a binomial distribution when the number of trials n gets large and the probability of success p is small. To see this, let $X \sim \mathcal{B}(n,p)$. Set $\lambda = np$. Then

$$\mathbb{P}(X=i) = \binom{n}{i}p^i(1-p)^{n-i} = \binom{n}{i}\frac{\lambda^i}{n^i}(1-\frac{\lambda}{n})^{n-i} = \frac{n(n-1)\dots(n-i+1)}{n^i}\,\frac{\lambda^i}{i!}(1-\frac{\lambda}{n})^{n-i}$$

For large n and keeping λ and i fixed, we see that

$$\frac{n(n-1)\dots(n-i+1)}{n^i} \approx 1, \quad (1-\frac{\lambda}{n})^{-i} \approx 1, \quad (1-\frac{\lambda}{n})^n \approx e^{-\lambda}.$$

Therefore, we have

$$\mathbb{P}(X=i) \approx e^{-\lambda}\frac{\lambda^i}{i!}.$$

Example 3.4. *A factory produces nails and packs them in boxes of 400. If the probability that a nail is substandard is 0.005, find the probability that a box selected at random contains at most two nails which are substandard.*

Solution: *Let X denotes the number of substandard nails in a box of 400. Then $X \sim \mathcal{B}(n,p)$ with $n = 400$ and $p = 0.005$. Since n is large enough and p is small, we can use Poisson* *approximation and therefore* $X \sim \mathcal{P}(\lambda)$ *where* $\lambda = 400 \times 0.005 = 2$. *The required probability is given by*

$$\mathbb{P}(X \le 2) = e^{-2}\big(1 + \frac{2^1}{1!} + \frac{2^2}{2!}\big) = 5\,e^{-2}\,.$$

Expected value and variance of Poisson random variable: In view of the relation between binomial distribution and Poisson distribution, one can expect that expectation and variance of a Poisson random variable would be $np = \lambda$. We now verify this result. For any $X \sim \mathcal{P}(\lambda)$,

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=0}^{\infty} ip(i) = \lambda e^{-\lambda}\sum_{i=1}^{\infty}\frac{\lambda^{i-1}}{(i-1)!} = \lambda\,,\\
\mathbb{E}[X^2] &= \sum_{i=0}^{\infty} i^2 p(i) = \lambda\sum_{i=1}^{\infty} ie^{-\lambda}\frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda}\sum_{j=0}^{\infty}(j+1)\frac{\lambda^j}{j!}\\
&= \lambda\Big(e^{-\lambda}\sum_{j=0}^{\infty}\frac{j\lambda^j}{j!} + e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}\Big) = \lambda(\lambda+1)\,,\\
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2 = \lambda(\lambda+1) - \lambda^2 = \lambda\,.
\end{aligned}$$

Moment generating function of X is given by $m_X(t) = e^{-\lambda(1-e^t)}$ for all t, see Example 2.16.

Example 3.5. *The mean number of errors due to a bug occurring in a minute is* 0.0001. *What is the probability that there will be no error in* 30 *minutes? How long would the program need to run to ensure that there will be a* 99.95% *chances that an error will show up to highlight this bug?*

Solution. *Let* X *denotes total number of errors occurring in* 30 *minutes. Then* $X \sim \mathcal{P}(\lambda)$ *where* $\lambda = 30 \times 0.0001 = 0.003$. *The probability that there will be no error in* 30 *minutes is equal to*

$$\mathbb{P}(X = 0) = e^{-0.003}.$$

Suppose that program need to run k *minutes . Then,* $\mathbb{P}\big(\text{no occurrence of error in } k \text{ minutes}\big) = e^{-(0.0001)k}$. *Hence, to be* 99.95% *sure catching the bug, we need to choose* k *such that*

$$1 - e^{-(0.0001)k} \ge 0.9995 \Leftrightarrow e^{-(0.0001)k} \le 0.0005 \implies k > 75000\,.$$

3.3. **Negative binomial and geometric random variables.** A positive, integer valued random variable X is said to be a **geometric random variable** with parameter p, denoted it by $X \sim \mathcal{G}(p)$, if its probability mass function is given by

$$\mathbb{P}(X = n) = (1-p)^{n-1}p \quad n = 1, 2, \ldots \text{ for some } 0 < p < 1.$$

X basically represents the number of trials required until a success occurs. One can easily check that

$$\mathbb{E}[X] = \frac{1}{p}, \quad \mathbb{E}[X^2] = \frac{2-p}{p^2}, \quad \mathrm{Var}(X) = \frac{1-p}{p^2}\,.$$

Example 3.6. *An urn contains* 10 *white and* 4 *black balls. Balls are randomly selected, one at a time, until a black one is obtained. If we assume that each ball selected is replaced before the next one is drawn, what is the probability that exactly* 3 *draws are needed ?*

Solution: *We denote by* X, *the number of draws required to select a black ball. Then* X *is a geometric random variable with parameter* $p = \frac{2}{7}$. *Hence the probability that exactly* 3 *draws are needed is given by*

$$\mathbb{P}(X = 3) = \big(\frac{5}{7}\big)^2 \times \frac{2}{7} = \frac{50}{243}\,.$$

Any random variable X whose probability mass function is given by

$$p(n) = \binom{n-1}{r-1} p^r (1-p)^{n-r} \quad n = r, r+1, \dots$$

is said to be a negative binomial random variable with parameters (r, p). It is denoted by $X \sim \mathcal{B}_N(r, p)$. It basically represents the number of trials required until a total of r success occurs. Note that a geometric random variable is just a negative binomial with parameter $(1, p)$. Let us calculate the mean and variance of negative binomial random variable X. Observe that

$$\begin{aligned}
\mathbb{E}[X^k] &= \sum_{n=r}^{\infty} n^k \binom{n-1}{r-1} p^r (1-p)^{n-r} = \frac{1}{p} \sum_{n=r}^{\infty} n^{k-1} n \binom{n-1}{r-1} p^{r+1} (1-p)^{n-r} \\
&= \frac{1}{p} \sum_{n=r}^{\infty} n^{k-1} r \binom{n}{r} p^{r+1} (1-p)^{n-r} = \frac{r}{p} \sum_{m=r+1}^{\infty} (m-1)^{k-1} \binom{m-1}{r} p^{r+1} (1-p)^{m-(r+1)} \qquad (3.3) \\
&= \frac{r}{p} \mathbb{E}\Big[(Y-1)^{k-1}\Big], \qquad (3.4)
\end{aligned}$$

where $Y \sim \mathcal{B}_N(1+r, p)$. Hence, setting $k = 1$, we have $\mathbb{E}[X] = \frac{r}{p}$. For $k = 2$, we have

$$\mathbb{E}[X^2] = \frac{r}{p} \mathbb{E}\Big[(Y-1)\Big] = \frac{r}{p} \big(\frac{1+r}{p} - 1 \big).$$

Using these estimate, we calculate the variance as

$$\mathrm{Var}(X) = \frac{r}{p} \big(\frac{1+r}{p} - 1 \big) - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}.$$

Example 3.7. *Suppose that Indian football team has ability to win any one game is 45% and games are independent of one another.*

a) What is the probability that it takes 15 games to win 4th game?
b) What is the expected value and variance of the number of games it will take to win their 40th game.
c) Knowing that they got their 45th win with 4 games remaining in the season, what is the probability that they do not get their 46th or more wins?

Solution: *Let X respectively Y be the number of games required to win their 4th respectively 40th game. Then $X \sim \mathcal{B}_N(4, 0.45)$ and $Y \sim \mathcal{B}_N(40, 0.45)$.*

a) The required probability is given by

$$\mathbb{P}(X = 15) = \binom{14}{3} (0.45)^4 (0.55)^{11}.$$

b) Expected value and variance of the number of game is given by

$$\mathbb{E}[Y] = \frac{40}{0.45} \approx 88.88, \quad \mathrm{Var}(Y) = \frac{40 \times 0.55}{(0.45)^2} \approx 108.6420.$$

c) Let Z denotes the number of games to get first win. Then $Z \sim \mathcal{G}(0.45)$. The required probability is given by

$$\mathbb{P}(Z > 4) = (1 - 0.45)^{4-1} = (0.55)^3 \approx 0.1664.$$

3.4. **Uniform random variable.** A random variable X is said to have a uniform distribution on the interval $[a,b]$ with $-\infty < a < b < \infty$, if its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise.} \end{cases}$$

We write $X \sim \mathcal{U}[a,b]$ if X has a uniform distribution on $[a,b]$. The distribution function of X is given by

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

One can easily check that

$$\mathbb{E}[X] = \frac{a+b}{2}, \quad \mathbb{E}[X^k] = \frac{b^{k+1} - a^{k+1}}{(k+1)b-a}, \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}\,.$$

Example 3.8. *Trains arrive at a specified station* **A** *at* 15*-minute intervals starting at* 8 *a.m. Suppose the passengers arrive at the station at a time that is uniformly distributed between* 8 *and* 8 : 30*. Find the probability that passenger waits* (i) *less than* 5 *minutes for a train, and* (ii) *more than* 10 *minutes for a train.*

Solution: *Let X denotes the time (in minutes) past* 8 *that the passengers arrive at the station* **A***. Then $X \sim \mathcal{U}[0,30]$. For less than* 5 *minutes waiting for a train, passengers need to arrive between* 8.10-8.15 *or between* 8.25-8.30*. Hence*

$$\begin{aligned} \mathbb{P}\Big(\textit{passengers wait less than } 5 \textit{ minutes}\Big) &= \mathbb{P}(10 < X < 15) + \mathbb{P}(25 < X < 30) \\ &= \int_{10}^{15} \frac{1}{30}\,dx + \int_{25}^{30} \frac{1}{30}\,dx = \frac{1}{3}\,. \end{aligned}$$

Similarly, for more than 10 minutes waiting, passengers need to arrive between 8-8.05 or between 8.15-8.20. Hence

$$\begin{aligned} \mathbb{P}\Big(\text{passengers wait more than 10 minutes}\Big) &= \mathbb{P}(0 < X < 5) + \mathbb{P}(15 < X < 20) \\ &= \int_{0}^{5} \frac{1}{30}\,dx + \int_{15}^{20} \frac{1}{30}\,dx = \frac{1}{3}\,. \end{aligned}$$

3.5. **Normal distribution.** The normal distribution is one of the most important distribution in the study of probability and mathematical statistics.

Definition 3.3. A random variable X is said to have a normal distribution with parameters μ and σ^2, denoted as $X \sim \mathcal{N}(\mu, \sigma^2)$, if its density function is given by

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

For $\mu = 0$ and $\sigma = 1$, the random variable $X \sim \mathcal{N}(0,1)$ is called standard normal random variable.

The distribution function of $X \sim \mathcal{N}(\mu, \sigma^2)$ is given by

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

Lemma 3.1. *For any $a, b \in \mathbb{R}$ with $a \neq 0$, define a random variable $Y := aX + b$, where $X \sim \mathcal{N}(\mu, \sigma^2)$. Then $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.*

Proof. For simplicity, let $a > 0$. Similar proof can be carrier out for $a < 0$. Let $F_Y(\cdot)$ be the cumulative distribution function of Y. Then

$$F_Y(x) = \mathbb{P}(Y \le x) = \mathbb{P}(X \le \frac{x-b}{a}) = F_X(\frac{x-b}{a}),$$

where $F_X(\cdot)$ is the cumulative distribution function of X. If f_Y is the density function of Y, then we have

$$f_Y(x) = \frac{1}{a} f_X(\frac{x-b}{a}) = \frac{1}{a\,\sigma\sqrt{2\pi}} e^{-\frac{(x-(a\mu+b))^2}{2a^2\sigma^2}}.$$

This implies that $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. □

Remark 3.2. In view of Lemma 3.1, we see that for any $X \sim \mathcal{N}(\mu, \sigma^2)$ the random variable $Z := \frac{X-\mu}{\sigma}$ is a standard normal random variable.

The moment generating function of $X \sim \mathcal{N}(\mu, \sigma^2)$ is given by

$$\begin{aligned} m_X(t) &:= \mathbb{E}[e^{tX}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = e^{\mu t + \frac{\sigma^2 t^2}{2}} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}}\,dx \\ &= e^{\mu t + \frac{\sigma^2 t^2}{2}}. \end{aligned}$$

One can easily deduce that

$$\mathbb{E}[X] = \mu, \quad \mathrm{Var}(X) = \mu^2.$$

Clearly, the central moments of odd order are all zero. Central moments of even order can be calculated as follows:

$$\begin{aligned} \mathbb{E}[(X-\mu)^{2n}] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{\infty} y^{2n} e^{-\frac{y^2}{2\sigma^2}}\,dy \\ &= \frac{\sigma^{2n}}{\sqrt{2\pi}} 2^{n+\frac{1}{2}} \int_0^{\infty} z^{n-\frac{1}{2}} e^{-z}\,dz = \frac{\sigma^{2n}}{\sqrt{2\pi}} 2^{n+\frac{1}{2}}\,\Gamma(n+\frac{1}{2}) \; \big(\therefore \Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t}\,dt\big) \\ &= \frac{\sigma^{2n}}{\sqrt{2\pi}} 2^{n+\frac{1}{2}} \frac{(2n-1)!}{2^n} \sqrt{\pi} = (2n-1)!\,\sigma^{2n}. \end{aligned}$$

It is customary to denote the cumulative distribution function of a standard normal random variable by $\Phi(x)$ i.e.,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy \quad -\infty < x < \infty.$$

For $x \ge 0$, the values of $\Phi(x)$ can be found from the Table 1. For negative values of x, $\Phi(x)$ is calculated from the relationship

$$\Phi(-x) = 1 - \Phi(x) \quad -\infty < x < \infty.$$

Example 3.9. *For any* $X \sim \mathcal{N}(3, 9)$, *find* i) $\mathbb{P}(2 < X < 5)$ *and* ii) $\mathbb{P}(|X-3| > 6)$.
Solution: *In view of Remark 3.2, we need to find the probability in terms of standard normal random variable and then use Table 1.*

i) $\mathbb{P}(2 < X < 5) = \mathbb{P}\Big(\frac{2-3}{3} < Z < \frac{5-3}{3}\Big) = \Phi(\frac{2}{3}) - \Phi(-\frac{1}{3}) = \Phi(\frac{2}{3}) + \Phi(\frac{1}{3}) - 1] \approx 0.3779.$

ii) $$\begin{aligned} \mathbb{P}(|X-3| > 6) &= 1 - \mathbb{P}(|X-3| \le 6) = 1 - \mathbb{P}(-3 \le X \le 9) = 1 - \mathbb{P}(-2 \le Z \le 2) \\ &= 1 - \big[\Phi(2) - \Phi(-2)\big] = 2(1 - \Phi(2)) \approx 0.456. \end{aligned}$$

TABLE 1. Area $\Phi(x)$ under the standard normal curve to the left of x

x	**0**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999
3.1	0.99903	0.99906	0.9991	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.9994	0.99942	0.99944	0.99946	0.99948	0.9995
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.9996	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.9997	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.9998	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.9999	0.9999	0.9999	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998

Example 3.10. *Suppose that the life lengths of two electronic devices say, D_1 and D_2, have normal distributions $\mathcal{N}(40, 36)$ and $\mathcal{N}(45, 9)$, respectively. If a device is to be used for 45 hours, which device would be preferred?*

Solution: *We will find which device has greater probability of lifetime more than 45 hours. We have*

$$\mathbb{P}(D_1 > 45) = \mathbb{P}(Z > \frac{5}{6}) = 1 - \Phi(\frac{5}{6}) \approx 0.2005$$
$$\mathbb{P}(D_2 > 45) = \mathbb{P}(Z > 0) = 1 - \Phi(0) = 0.5$$

Since $\mathbb{P}(D_2 > 45) > \mathbb{P}(D_1 > 45)$, the device D_2 will be preferred.

3.5.1. **The Normal Approximation to the Binomial Distribution.** : We have seen that $\mathcal{B}(n,p)$ converges to $\mathcal{P}(\lambda = np)$ when $n$is large and p is very small. One can ask the limiting distribution of $\mathcal{B}(n.p)$ when neither p or $q = 1-p$ is very small (in particular when $np(1-p)$ is large). This result is known as **DeMoivre-Laplace limit theorem**. We state this theorem without proof as it as a special case of central limit theorem which will be discussed later.

Theorem 3.2 (DeMoivre-Laplace limit theorem)**.** *If S_n denotes the number of successes, with p as a success of probability, that occur when n independent trials are performed, then for any $a < b$,*

$$\mathbb{P}\Big(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\Big) \to \Phi(b) - \Phi(a) \quad \text{as } n \to \infty.$$

Note that the normal approximation will, in general, be quite good for values of n satisfying $np(1-p) \ge 10$.

Example 3.11. *A die is tossed 1000 consecutive times. Calculate the probability that the number 4 shows up between 150 and 200 times. What is the probability that the number 4 appears exactly 150 times?*
Solution: *If X denotes the number of times 4 shows up then $X \sim \mathcal{B}(1000, \frac{1}{6})$. Since $np(1-p) = \frac{5000}{36} > 10$, we can use normal approximation of binomial distribution.*

$$\mathbb{P}(150 \le X \le 200) = \mathbb{P}\Big(\frac{150 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{X - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{200 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}}\Big)$$
$$= \mathbb{P}\big(-1.14142 \le Z \le 2.8284\big) = \Phi(2.8284) - \Phi(-1.14142) \approx 0.9183$$

Since binomial is a discrete integer-valued random variable, whereas the normal is a continuous random variable,it is best to write $\mathbb{P}(X = i)$ as $\mathbb{P}(i - \frac{1}{2} \le X \le i + \frac{1}{2})$ before applying the normal approximation (this is called the continuity correction). Therefore, we have

$$\mathbb{P}(X = 150) = \mathbb{P}(149.5 < X < 150.5) = \mathbb{P}\Big(\frac{149.5 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{X - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{150.5 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}}\Big)$$
$$= \mathbb{P}\big(-1.4566 \le Z \le -1.3718\big) = \Phi(-1.3718) - \Phi(-1.4566) \approx 0.01319\,.$$

3.6. **Exponential random variable.** A continuous random variable is said to be exponential random variable with parameter $\lambda > 0$, denoted by $\mathrm{Exp}(\lambda)$, if its pdf is given by

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

The cdf of an exponential random variable is given by

$$F_X(x) = \mathbb{P}(X \le x) = \int_0^x f(y)\,dy = 1 - e^{-\lambda x}, \quad x \ge 0.$$

Mean and variance of $X \sim \mathrm{Exp}(\lambda)$: By using integration by parts formula, we have, for $n > 0$

$$
\begin{aligned}
\mathbb{E}[X^n] &= \int_0^\infty x^n \lambda e^{-\lambda x}\,dx = -\int_0^\infty x^n \frac{d}{dx} e^{-\lambda x} = n\int_0^\infty e^{-\lambda x} x^{n-1}\,dx \\
&= \frac{n}{\lambda}\int_0^\infty x^{n-1}\lambda e^{-\lambda x}\,dx = \frac{n}{\lambda}\mathbb{E}[X^{n-1}] \\
&\implies \mathbb{E}[X] = \frac{1}{\lambda}, \quad \mathbb{E}[X^2] = \frac{2}{\lambda}\mathbb{E}[X] = \frac{2}{\lambda^2}\,.
\end{aligned}
$$

Hence, we have

$$\mathrm{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}\,.$$

In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs e.g., the amount of time until an earthquake occurs, a new war breaks out etc.

Definition 3.4. We say that a non-negative random variable X is **memoryless** if

$$\mathbb{P}\Big(X > s+t \big| X > t\Big) = \mathbb{P}(X > s) \Leftrightarrow \mathbb{P}(X > t+s) = \mathbb{P}(X > t)\mathbb{P}(X > s) \quad \forall\ s,t \geq 0. \tag{3.5}$$

Exponentially distributed random variables are memoryless. Indeed, for any $t, s \geq 0$, one has

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

Moreover, we have the following theorem.

Theorem 3.3. *Let X be a random variable such that $\mathbb{P}(X > 0) > 0$. Then $X \sim \mathrm{Exp}(\lambda)$ if and only if (3.5) holds.*

Example 3.12. *Suppose the time of use it takes until a smartphone fails is modelled by an exponential random variable. The average time until failure is 1000 hours. What is the probability that the smartphone*

i) fails in the first 10 hours?
ii) does not fail in the first 1000 hours?
iii) does not fail in the next 500 hours knowing that it already used for 200 hours?

Solution: *Let X be time in hours until failure. Then $X \sim \mathrm{Exp}(\lambda)$ for some $\lambda > 0$. Given that $\mathbb{E}[X] = 1000$. Hence $\lambda = \frac{1}{1000}$.*

i) $\mathbb{P}(X < 10) = \displaystyle\int_0^{1000} \frac{1}{1000} e^{-\frac{x}{1000}}\,dx = 1 - e^{-\frac{1}{100}}$.
ii) $\mathbb{P}(X \geq 1000) = 1 - \mathbb{P}(X \leq 1000) = e^{-1} \approx 0.3678$.
iii) By memoryless property, the required probability is

$$\mathbb{P}\Big(X > 500 + 200 \big| X > 200\Big) = \mathbb{P}(X > 500) = e^{-\frac{1}{2}}\,.$$

3.6.1. **Failure rate function:** Let X be a positive continuous random variable, interpreted as a lifetime of some item, with distribution function $F(\cdot)$ and density function $f(\cdot)$. The failure rate function $\lambda(t)$ of F is defined by

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}, \quad \text{where } \bar{F} = 1 - F.$$

$\lambda(t)$ can be represented as the conditional probability intensity that a t-unit-old item will fail i.e.,

$$\mathbb{P}\Big(X \in (t, t+dt) \big| X > t\Big) \approx \frac{f(t)}{\bar{F}(t)}\,dt.$$

Reliability is defined as the probability that a system (component) will function over some time period t. The function $\bar{F}(t)$ is known as **reliability function**. Sometimes it is denoted by $R(t) := \mathbb{P}(X > t)$.

One can easily check that for an exponentially distributed random variable, the failure rate function $\lambda(t)$ is constant and equal to λ. The failure rate function $\lambda(t)$ uniquely determines the distribution F. Indeed, since $\lambda(t) = \dfrac{\frac{d}{dt}F(t)}{1-F(t)}$, we have

$$\log(1-F(t)) = -\int_0^t \lambda(s)\,ds + c.$$

Since X is positive random variable, one has $F(0) = 0$ and hence $c = 0$. Thus,

$$F(t) = 1 - \exp\Big\{-\int_0^t \lambda(s)\,ds\Big\}.$$

For instance, if $\lambda(t) = a + bt$, then the distribution function and density function of the random variable is given by

$$F(t) = 1 - e^{-(at+\frac{bt^2}{2})}, \quad f(t) = (a+bt)e^{-(at+\frac{bt^2}{2})}, \quad t \ge 0\,.$$

A **design life** is defined to be the time to failure t_R that corresponds to a specified reliability R i.e., $R(t_R) = R$.

Example 3.13. *Given the failure rate function $\lambda(t) = 5\times 10^{-6}t$ where t is measured in operating hours, what is the design life if a 0.98 reliability is desired?*
Solution: *The reliability function can be written in terms of failure rate function as $R(t) = \exp\{-\int_0^t \lambda(s)\,ds\}$. By the given condition, we have*

$$0.98 = \exp\{-\int_0^t 5\times 10^{-6} s\,ds\} \implies t = \sqrt{-\frac{\ln(0.98)}{2.5\times 10^{-6}}} = 89.89 \approx 90\,.$$

Therefore the design life is 90 hours.

3.7. **Gamma distribution:** The gamma distribution is used to describe the intervals of time between two consecutive failures of an airplane's motor or the intervals of time between arrivals of clients to a queue in a supermarket's cashier point. A random variable X is said to have a gamma distribution with positive parameters (α, λ), if its density function is given by

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & x \ge 0 \\ 0, \quad x < 0 \end{cases}$$

where $\Gamma(\alpha)$ is the gamma function. The parameter α is called the shape parameter while λ is the scale parameter. We denote gamma distribution as $\Gamma(\alpha, \lambda)$.

The distribution function of $X \sim \Gamma(\alpha, \lambda)$ is given by

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

The moment generating function of $X \sim \Gamma(\alpha, \lambda)$ is given by

$$m_X(t) = \Big(\frac{\lambda}{\lambda - t}\Big)^{\alpha} \text{ if } t < \alpha.$$

Further,

$$\mathbb{E}[X] = \frac{d}{dt}m_X(t)\Big|_{t=0} = \frac{\alpha}{\lambda}, \quad \mathbb{E}[X^2] = \frac{d^2}{dt^2}m_X(t)\Big|_{t=0} = \frac{\alpha^2+\alpha}{\lambda^2}, \quad \mathrm{Var}(X) = \frac{\alpha}{\lambda^2}\,.$$

- Gamma distribution with parameters $(1, \lambda)$ is nothing but the exponential distribution with parameter λ.
- Gamma distribution with parameters $(\frac{k}{2}, \frac{1}{2})$ with $k \in \mathbb{N}$ is called **chi-square** distribution with k degrees of freedom. We denote the chi-square distribution by $X \sim \chi^2_{(k)}$.
- For $\alpha > 1$ and $\lambda > 0$, the gamma distribution $\Gamma(\alpha, \lambda)$ is called **Erlang distribution**.

Example 3.14. *Time spent on a computer (X) is gamma distributed with mean 20 min and variance 80 min^2. Find $\mathbb{P}(X < 24)$ and $\mathbb{P}(20 < X < 40)$.*
Solution: $X \sim \Gamma(\alpha, \lambda)$. *Given that $\mathbb{E}[X] = 20$ and $\mathrm{Var}(X) = 80$ i.e.,*

$$\frac{\alpha}{\lambda} = 20, \quad \frac{\alpha}{\lambda^2} = 80.$$

Solving these, we have $\alpha = 5$ and $\lambda = \frac{1}{4}$. Therefore,

$$\mathbb{P}(X < 24) = \frac{1}{\Gamma(5)} \int_0^6 y^{\alpha-1} e^{-y}\, dy \approx 0.715\,,$$
$$\mathbb{P}(20 < X < 40) = \frac{1}{\Gamma(5)} \int_5^{10} y^{\alpha-1} e^{-y}\, dy \approx 0.411\,.$$

3.8. **Weibull distribution:** The Weibull distribution is used extensively in reliability and life data analysis such as situations involving failure times of items. Since this failure time may be any positive number, the distribution is continuous. It has been used successfully to model such things as vacuum tube failures and ball bearing failures. A random variable X ha Weibull distribution with parameters (γ, μ, α) if its density function is given by

$$f(x) = \begin{cases} \frac{\gamma}{\alpha}\Big(\frac{x-\mu}{\alpha}\Big)^{\gamma-1} \exp\Big(-\big(\frac{x-\mu}{\alpha}\big)^{\gamma}\Big), & x \ge \mu \\ 0, \quad x < \mu\,, \end{cases}$$

where γ, α are positive constant. We denote the Weibull distribution as $W(\gamma, \mu, \alpha)$. The value γ is called the shape parameter, μ is the location parameter and α is the scale parameter.

Example 3.15. *The time to failure (in hours) of bearings in a mechanical shaft is satisfactorily modelled as a Weibull random variable with $\gamma = 0.5$, $\mu = 0$ and $\alpha = 5000$. Determine the probability that a bearing lasts fewer than 6000 hours and also the mean time to failure.*
Solution: *If X denotes the time (in hours) to failure of bearings then $X \sim W(0.5, 0, 5000)$. Observe that the cumulative distribution function $F(x)$ is given by*

$$F(x) = 1 - e^{-\big(\frac{x}{5000}\big)^{\frac{1}{2}}}, \quad x \ge 0.$$

The probability that a bearing lasts fewer than 6000 hours is given by

$$\mathbb{P}(X < 6000) = F(6000) = 1 - e^{-\big(\frac{6}{5}\big)^{0.5}} \approx 0.666\,.$$

The expected value of X is

$$\mathbb{E}[X] = \frac{1}{2}\int_0^\infty \big(\frac{x}{\alpha}\big)^{0.5} e^{-\big(\frac{x}{\alpha}\big)^{0.5}}\, dx = \frac{\alpha}{2}\int_0^\infty \sqrt{y} e^{-\sqrt{y}}\, dy = \alpha \int_0^\infty y^2 e^{-y}\, dy = \alpha\Gamma(3) = 2\alpha$$

Thus, the mean time to failure of bearing is 1000 hours.

3.9. **Beta distribution.** A random variable is said to have a beta distribution with parameter $a > 0$ and $b > 0$, denoted as $\beta(a, b)$, if its density function id given by

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}\chi_{(0,1)}(x)$$

where $B(a,b)$ is the beta function i.e., $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$. The beta distribution can be used to model random phenomenon whose set of possible values is some finite interval. One can easily check that

$$\mathbb{E}[X] = \frac{a}{a+b}, \quad \mathrm{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

Example 3.16. *During an 8-hour shift, the portion of time Y that a machine is down for maintenance or repairs has a beta distribution with $a = 1$ and $b = 2$. The cost of this downtime, due to lost production and cost of maintenance and repair is given by $C = 5 + 20Y + 5Y^2$. Find the mean of C.*

Solution: *Given that $Y \sim \beta(1,2)$ and hence , we have*

$$\mathbb{E}[Y] = \frac{1}{3}, \quad \mathrm{Var}(Y) = \frac{2}{9 \times 4} = \frac{1}{18} \implies \mathbb{E}[Y^2] = \mathrm{Var}(Y) + \big(\mathbb{E}[Y]\big)^2 = \frac{1}{6}.$$

Thus, the mean of C is given by

$$\mathbb{E}[C] = 5 + 20\mathbb{E}[Y] + 5\mathbb{E}[Y^2] = \frac{75}{6}.$$

3.10. **Cauchy distribution.** A random variable X is said to have a Cauchy distribution with parameter θ, $-\infty < \theta < \infty$, if its density function is given by

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2} \qquad -\infty < x < \infty.$$

Its distribution function is given by

$$F_X(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+(y-\theta)^2}\,dy = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x-\theta).$$

One can easily check that $\mathbb{E}[X^r]$ does not exist for any $r \ge 1$. The characteristic function of X is given by

$$\phi_X(t) = \exp(i\theta t - |t|) \quad \forall\, t \in \mathbb{R}.$$

A Cauchy distribution with parameter $\theta = 0$ is known as standard Cauchy distribution.

Example 3.17. *Fine the distribution of $X := \tan(Y)$ for any $Y \sim \mathcal{U}(\frac{\pi}{2}, \frac{\pi}{2})$.*

Solution: *The distribution function of X is given by*

$$F_X(x) = \mathbb{P}(\tan(Y) \le x) = \mathbb{P}(Y \le \tan^{-1}(x)) = \frac{\tan^{-1}(x) + \frac{\pi}{2}}{\pi} = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x).$$

Hence the density function of X is given by

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{1}{\pi}\frac{1}{1+x^2} \qquad -\infty < x < \infty.$$

Thus, X has a Cauchy distribution with parameter $\theta = 0$.