Prepared by: Dr. Haifa Abdul-Jawad & Dr. Zaid T. Al-Khaledi

The beta distribution: A continuous random variable X is said to have a *beta distribution*, denoted by $X \sim Beta(\alpha, \beta)$, if its PDF has the following form:

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

where $\alpha > 0$ is shape 1 parameter, and $\beta > 0$ is shape 2 parameter. The quantity $\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is called the *Complete Beta function*, denoted by $B(\alpha, \beta)$. That is

$$\begin{aligned} B(\alpha, \beta) &= \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!} \end{aligned}$$

The beta distribution is a suitable model for the random behavior of percentages and proportions. If $\alpha = \beta = 1$, then the beta distribution reduces to the continuous uniform distribution.

Mean, variance, and the r-th mean: By definition of the mean, the variance, and the r-th are defined as follows:

$$\begin{aligned}
E[X] &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x\left(x^{\alpha-1}\right)(1-x)^{\beta-1}dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+1)-1}(1-x)^{\beta-1}dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}\right), \qquad \text{by the definition of beta function} \\
&= \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!} \left(\frac{(\alpha)!(\beta-1)!}{(\alpha+\beta)!}\right) = \frac{\alpha}{\alpha+\beta}
\end{aligned}$$

$$\begin{aligned}
E[X^2] &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^2\left(x^{\alpha-1}\right)(1-x)^{\beta-1}dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+2)-1}(1-x)^{\beta-1}dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)}\right), \qquad \text{by the definition of beta function} \\
&= \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!} \left(\frac{(\alpha+1)!(\beta-1)!}{(\alpha+\beta+1)!}\right) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(X) &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} \\
&= \frac{\alpha}{(\alpha+\beta)} \left(\frac{(\alpha+1)}{(\alpha+\beta+1)} - \frac{\alpha}{(\alpha+\beta)}\right) \\
&= \frac{\alpha}{(\alpha+\beta)} \left(\frac{(\alpha+1)(\alpha+\beta) - \alpha(\alpha+\beta+1)}{(\alpha+\beta)(\alpha+\beta+1)}\right) \\
&= \frac{\alpha}{(\alpha+\beta)} \left(\frac{\cancel{\alpha^2} + \cancel{\alpha\beta} + \cancel{\alpha} + \beta - \cancel{\alpha^2} - \cancel{\alpha\beta} - \cancel{\alpha}}{(\alpha+\beta)(\alpha+\beta+1)}\right) \\
&= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{aligned}$$

$$
\begin{aligned}
E[X^r] &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^r \left(x^{\alpha-1}\right)(1-x)^{\beta-1} dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+r)-1}(1-x)^{\beta-1} dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{\Gamma(\alpha+r)\Gamma(\beta)}{\Gamma(\alpha+\beta+r)}\right), \qquad \text{by the definition of beta function} \\
&= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+r)}{\Gamma(\alpha)\Gamma(\alpha+\beta+r)}
\end{aligned}
$$

Example 6.1: Let $f(x) = 12x^2(1-x)$, for $0 \leq x \leq 1$ be the PDF of X. Calculate the mean and the variance of X.

Solution: Notice that X has a beta distribution with $\alpha = 3$ and $\beta = 2$ (i.e. $X \sim B(3,2)$). Therefore

$$
\begin{aligned}
E[X] &= \frac{\alpha}{\alpha+\beta} \\
&= \frac{3}{3+2} = \frac{3}{5}
\end{aligned}
$$

$$
\begin{aligned}
\text{Var}(X) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{3(2)}{(3+2)^2(3+2+1)} \\
&= \frac{6}{25(6)} = \frac{1}{25}
\end{aligned}
$$

Example 6.2: Let $X \sim Beta(2,2)$. Write down the PDF of X. Then calculate the mean and the variance of X.

Solution:

$$\begin{aligned} f(x) &= \frac{\Gamma(2+2)}{\Gamma(2)\Gamma(2)} x^{2-1}(1-x)^{2-1} \\ &= \frac{3!}{1!(1!)} x(1-x) = 6x(1-x) \end{aligned}$$

$$f(x) = \begin{cases} 6x(1-x) & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} E[X] &= \frac{\alpha}{\alpha+\beta} \\ &= \frac{2}{2+2} = \frac{1}{2} \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= \frac{2(2)}{(2+2)^2(2+2+1)} \\ &= \frac{4}{16(5)} = \frac{1}{20} \end{aligned}$$

Example 6.3 (Homework): Find the value of c in each of the following PDF's. Then find the mean and the variance of X:

1- $f(x) = cx(1-x)^3$, for $0 \le x \le 1$.

2- $f(x) = cx^4(1-x)^6$, for $0 \le x \le 1$.

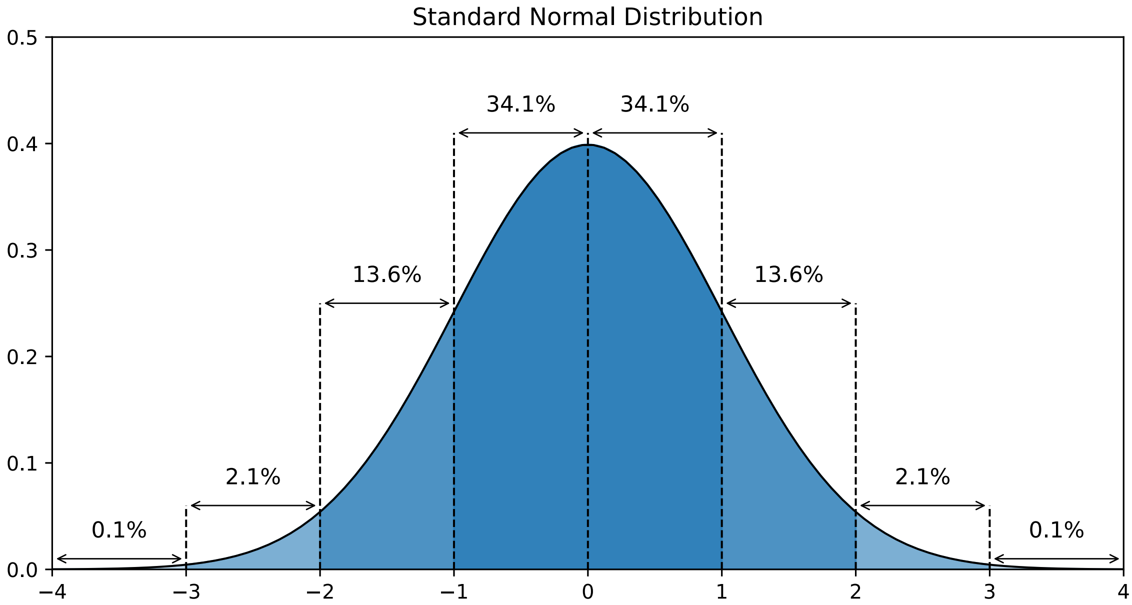
The normal distribution: Also known as the *Gaussian distribution.*

A continuous random variable X is said to have a normal distribution, denoted by $X \sim N(\mu, \sigma^2)$, if its PDF has the following form:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \text{ for } x \in \mathbb{R}$$

where $\mu \in \mathbb{R}$ is the mean of the distribution (the location parameter), and $\sigma^2 > 0$ is the variance of the distribution (σ is the scale parameter). If $\mu = 0$ and $\sigma^2 = 1$, then the distribution is called the *standard normal distribution*.

The normal distribution is the most important probability distribution in statistics because it accurately describes the distribution of values for many natural phenomena, like height, weight, age, IQ score, etc. The curve of the normal distribution has a bell shape. It is symmetric about the mean and indicates that values near the mean occur more frequently than values farther away from the mean. The mean, the median, and the mode of normal distribution are all equal.



Mean, variance, MGF, and CDF: By definition of the mean, the variance, the MGF, and the CDF are defined as follows:

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \end{aligned}$$

let $z = \frac{x-\mu}{\sigma}$, then $x = \sigma z + \mu$, and $dx = \sigma dz$. Hence

$$\begin{aligned} E[X] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (\sigma z + \mu) e^{-\frac{1}{2}z^2} \sigma dz \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \end{aligned}$$

notice that $\int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = -e^{-\frac{1}{2}z^2}$. And $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ is the PDF of $N(0,1)$.

Therefore $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 1$. Hence

$$\begin{aligned} E[X] &= \frac{\sigma}{\sqrt{2\pi}} \left(-\left(e^{-\frac{1}{2}(\infty)^2} - e^{-\frac{1}{2}(-\infty)^2} \right) \right) + \mu \\ &= 0 + \mu = \mu \end{aligned}$$

similarly, we can show that

$$E[X^2] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \sigma^2 + \mu^2$$

$$\begin{aligned} \mathrm{Var}(X) &= E[X^2] - E[X]^2 \\ &= \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \end{aligned}$$

$$\begin{aligned} M_X(t) &= E[e^{tX}] \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \end{aligned}$$

let $z = \frac{x-\mu}{\sigma}$, then $x = \sigma z + \mu$, and $dx = \sigma dz$. Hence

$$\begin{aligned} M_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma z+\mu)} e^{-\frac{1}{2}z^2} dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma z+\mu)} e^{-\frac{1}{2}z^2} dz \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(z^2-2t\sigma z\right)} dz \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(z^2-2t\sigma z+t^2\sigma^2-t^2\sigma^2\right)} dz \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu} e^{\frac{1}{2}t^2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t\sigma)^2} dz \end{aligned}$$

notice that $e^{-\frac{1}{2}(z-t\sigma)^2}$ is the kernel of $N(t\sigma, 1)$. Therefore

$$\begin{aligned} M_X(t) &= \frac{1}{\sqrt{2\pi}} e^{t\mu} e^{\frac{1}{2}t^2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t\sigma)^2} dz \\ &= \frac{1}{\sqrt{2\pi}} e^{t\mu+\frac{1}{2}t^2\sigma^2} \sqrt{2\pi} = e^{t\mu+\frac{1}{2}t^2\sigma^2} \end{aligned}$$

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\ &= P(Z \leq z) \\ &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du = \Phi(z) \end{aligned}$$

where $\Phi(z)$ is the cumulative distribution function of the standard normal distribution. The values of $\Phi(z)$ can be obtained from the z-tables or using any statistical software. Notice that if X has a normal distribution, then

$P(X \leq x) = P(X \geq -x)$ because the distribution is symmetric.

Example 6.4 (Homework): Let $X \sim N(\mu, \sigma^2)$. Show that the median and the mode of X equal to μ.

Example 6.5: Let X has the following PDF:

$$f(x) = ce^{-2x^2+8x}, \text{for } x \in \mathbb{R}$$

find the value of c.

Solution: Notice that

$$\begin{aligned}
f(x) &= ce^{-2x^2+8x} \\
&= ce^{-2(x^2-4x)} \\
&= ce^{-2(x^2-4x+2^2-2^2)} \\
&= ce^{-2(x-2)^2+8} \\
&= ce^8 e^{-\frac{4}{2}(x-2)^2} \\
&= ce^8 e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2}
\end{aligned}$$

Now we have

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= 1 \\
\Longrightarrow \quad ce^8 \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2} dx &= 1 \\
\Longrightarrow \quad ce^8 \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2} dx &= 1
\end{aligned}$$

notice that $e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2}$ is the kernel of normal distribution with $\mu = 2$ and $\sigma^2 = 1/4$. Therefor

$$\begin{aligned} ce^8 \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2} dx = 1 \\ \implies ce^8 \sqrt{2\pi}(1/2) = 1 \\ \implies c = \frac{1}{e^8 \sqrt{2\pi}(1/2)} \end{aligned}$$

hence

$$\begin{aligned} f(x) &= ce^{-2x^2+8x} \\ &= ce^8 e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2} \\ &= \frac{1}{e^8\sqrt{2\pi}(1/2)} e^8 e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2} \\ &= \frac{1}{\sqrt{2\pi}(1/2)} e^{-\frac{1}{2}\left(\frac{x-2}{1/2}\right)^2}, \text{for } x \in \mathbb{R} \end{aligned}$$