**Beta** 10M

# Beta

Let $X$ follow a *beta* distribution with parameters $a$, $b$, and $\theta$, i.e.

$$X \sim \text{Beta}\,(a,\, b,\, \theta)$$

Then, $X$ has the following PDF:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{x}{\theta}\right)^a \left(1 - \frac{x}{\theta}\right)^{b-1} \frac{1}{x}, \qquad 0 < x < \theta$$

A beta distribution's PDF is recognizable from its finite support and its purely polynomial terms, i.e. no negative powers of $x$, and no exponential or logarithmic terms.

# Coach's Remarks

Many students forget the $\frac{1}{x}$ term at the end even though the PDF is available on the exam table. Don't be one of them.

The mean and variance are

$$\mathrm{E}[X] = \frac{a}{a+b} \cdot \theta$$

$$\mathrm{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)} \cdot \theta^2$$

## Coach's Remarks

Most questions that use the beta distribution will set $\theta$ to $1$. In that case, the PDF simplifies to

$$f(x) = c \cdot x^{a-1}(1-x)^{b-1}, \qquad 0 < x < 1$$

which makes it easier to integrate and recognize as beta.

## Example S2.2.5.1

A random variable $X$ has the following PDF:

$$f(x) = 105x^4(1-x)^2, \qquad 0 < x < 1$$

Determine the second moment of $X$.

### Solution

Because the PDF has finite support and only consists of polynomial terms, it belongs to a beta distribution. The support is from 0 to 1, which means $\theta = 1$.

The PDF of a beta distribution with $\theta = 1$ is in the form of $c \cdot x^{a-1}(1-x)^{b-1}$. Compare this to the PDF given to deduce $a = 5$ and $b = 3$.

$$X \sim \text{Beta}\,(5,\,3,\,1)$$

Look up beta's moment formula in the exam table. The second moment is

$$\begin{aligned}\mathrm{E}\left[X^2\right] &= \frac{\theta^2 a\,(a+1)}{(a+b)\,(a+b+1)} \\ &= \frac{1^2\,(5)\,(5+1)}{(5+3)\,(5+3+1)} \\ &= \boldsymbol{\frac{5}{12}}\end{aligned}$$

## Alternative Solution

Alternatively, you can solve using first principles.

$$\begin{aligned}\mathrm{E}\left[X^2\right] &= \int_{-\infty}^{\infty} x^2 \cdot f(x)\,\mathrm{d}x \\ &= \int_0^1 x^2 \cdot 105\,x^4 (1-x)^2\,\mathrm{d}x \\ &= 105 \cdot \int_0^1 x^6\left(1-2x+x^2\right)\mathrm{d}x \\ &= 105 \cdot \int_0^1 \left(x^6 - 2x^7 + x^8\right)\mathrm{d}x \\ &= 105 \cdot \left[\frac{x^7}{7} - \frac{2x^8}{8} + \frac{x^9}{9}\right]_0^1 \\ &= \boldsymbol{\frac{5}{12}}\end{aligned}$$

## Uniform

A ***uniform*** distribution has a constant density. Let $X$ follow a uniform distribution on the interval $[a,\, b]$, i.e.

$$X \sim \text{Uniform}\,(a,\, b)$$

Then, $X$ has the following PDF:

$$f(x) = \frac{1}{b-a}, \qquad a \le x \le b$$

The mean is the **midpoint** of the interval.

$$\mathrm{E}[X] = \frac{a+b}{2}$$

The variance is

$$\mathrm{Var}[X] = \frac{(a-b)^2}{12}$$

The second raw moment can be easily calculated by adding the variance to the square of the mean. However, some students prefer this shortcut:

$$\mathrm{E}\left[X^2\right] = \frac{a^2+ab+b^2}{3}$$

## Coach's Remarks

A uniform distribution on the interval $[0,\ \theta]$ is equivalent to a beta distribution with parameters $a = b = 1$ and $\theta$.

Assume $X$ is uniformly distributed on the interval $[a,\ b]$. Then, $X$ given it is greater than $d$, where $a < d < b$, will be uniformly distributed on the interval $[d,\ b]$.

$$X \sim \text{Uniform}\ (a,\ b)$$

$$X \mid X > d \sim \text{Uniform}\ (d,\ b)$$

Shifting $X$ leftwards by $d$ will shift the endpoints of the interval by the same amount. Therefore,

$$X - d \mid X > d \sim \text{Uniform}\ (0,\ b - d)$$

The figure below illustrates the transition from $X$ to $X \mid X > d$ to $X - d \mid X > d$.

$f(x)$