MATH7501: Mathematics for Data Science I

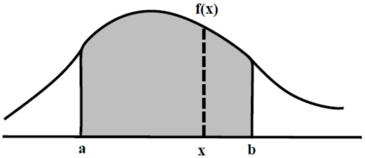
Unit 9.7: Introduction to Probability Theory

Probability Density Functions

Definition (Probability Density Function)

A random variable X is said to be *continuous* if there is a (piece-wise) continuous positive function $f(x)$ (called the *probability density function*) so that

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx \text{ for each } a \leq b.$$



Remarks

If f is a pdf, then by definition, $\int_{-\infty}^{\infty} f(x)dx = 1$. Also, $\mathbb{P}(X = a) = \int_a^a f(x)dx = 0$ for any a.

Probability Density Functions

Question (Drawing from an Interval)

Draw a number from the interval of real numbers $[0, 2]$ uniformly at random. If X is the real number selected, what is the pdf f?

Solution

The function will be

$$f(x) = \begin{cases} \frac{1}{2} \text{ if } 0 \leq x \leq 2, \\ 0 \text{ otherwise.} \end{cases}$$

Note that $\int_{-\infty}^{\infty} f(x)dx = \int_0^2 f(x)dx = 1$.

Cumulative Distribution Function

Definition (Cumulative Distribution Function)

If X is a continuous random variable, then the *cumulative distribution function* (cdf) is the function $F : \mathbb{R} \to [0, 1]$ with $F(x) = \mathbb{P}(X \leq x)$.

Properties of the cdf

If F is the cdf of a random variable X, then

- F is increasing, i.e., $F(x) \leq F(y)$ if $x \leq y$.
- $\lim_{x\to\infty} F(x) = 1$, and $\lim_{x\to-\infty} F(x) = 0$.
- The cdf F is continuous, and satisfies $F(x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u$. Therefore, $f(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x)$.

Cumulative Distribution Function

Question (Drawing from an Interval)

Draw a number from the interval of real numbers $[0, 2]$ uniformly at random. Find the cdf F, hence find the pdf f.

Solution

We have that

$$F(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \frac{x}{2}, & \text{if } 0 \leq x \leq 2, \\ 1, & \text{if } x \geq 2. \end{cases}$$

By differentiating, we find

$$f(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \frac{1}{2}, & \text{if } 0 \leq x \leq 2, \\ 0, & \text{if } x \geq 2. \end{cases}$$

Cumulative Distribution Function

Question (Drawing from an Interval)

Draw a number from the interval of real numbers $[0, 2]$ uniformly at random. Find the cdf F, hence find the pdf f.

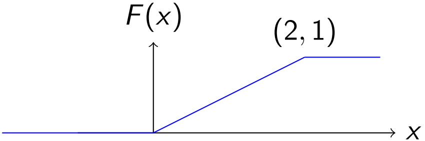


Figure: cdf

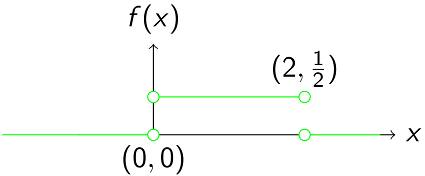


Figure: pdf

Expectation and Variance

Definition (Expectation for continuous Random Variables)

Suppose X is a continuous random variable with pdf f. For any real-valued function g, the expectation of $g(X)$ is

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Expectation and Variance

Question (Signals)

Let $Y = a\ cos(\omega t + X)$ be the value of a sinusoid signal at time t with uniform random phase $X \in (0, 2\pi]$. Find the expected value of the signal $\mathbb{E}[Y]$, and the expected power of the signal $\mathbb{E}[Y^2]$.

Solution

The pdf of X is $f(x) = \begin{cases} \frac{1}{2\pi} \text{ if } x \in [0, 2\pi), \\ 0 \text{ otherwise.} \end{cases}$ We therefore have $\mathbb{E}[Y] = \int_0^{2\pi} a\cos(\omega t + x)\frac{1}{2\pi}dx = 0$. Also,

$$\begin{aligned}
\mathbb{E}[Y^2] &= \int_0^{2\pi} a^2\cos^2(\omega t + x)\frac{1}{2\pi}dx \\
&= \int_0^{2\pi} a^2\left(\frac{1}{2} + \frac{1}{2}\cos(2\omega t + 2x)\right)\frac{1}{2\pi}dx = \frac{a^2}{2}.
\end{aligned}$$

Definition (Variance)

The *variance* of a random variable X is given by

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

and measures how far spread out the outcomes are from the expected value. The variance is sometimes denoted σ_X^2; the number σ_X is the *standard deviation* of X.

Expectation and Variance

Question (Properties of the Expectation)

Let X be a random variable, a, b real numbers, and g, h real-valued functions. Show that

- $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$, and
- $\mathbb{E}[g(X) + h(X)] = \mathbb{E}[g(X)] + \mathbb{E}[h(X)]$.

Question (Properties of the Variance)

Let X be a random variable, and a, b real numbers. Show that

- $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$, and
- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

Expectation and Variance

Solution (Properties of Expectation)

If X is continuous with pdf f, then

$$\mathbb{E}[aX + b] = \int_{-\infty}^{\infty} (ax + b) f(x) dx$$

and

Expectation and Variance

Solution (Properties of Expectation)

If X is continuous with pdf f, then

$$\begin{aligned}\mathbb{E}[aX + b] &= \int_{-\infty}^{\infty} (ax + b) f(x) dx \\ &= a \int_{-\infty}^{\infty} x f(x) dx + b = a\mathbb{E}[X] + b,\end{aligned}$$

and

Expectation and Variance

Solution (Properties of Expectation)

If X is continuous with pdf f, then

$$\begin{aligned}\mathbb{E}[aX+b] &= \int_{-\infty}^{\infty}(ax+b)f(x)dx \\ &= a\int_{-\infty}^{\infty}xf(x)dx + b = a\mathbb{E}[X]+b,\end{aligned}$$

and

$$\mathbb{E}[g(X)+h(X)] = \int_{-\infty}^{\infty}\left(g(x)+h(x)\right)f(x)dx$$

Expectation and Variance

Solution (Properties of Expectation)

If X is continuous with pdf f, then

$$\begin{aligned}\mathbb{E}[aX+b] &= \int_{-\infty}^{\infty}(ax+b)f(x)dx \\ &= a\int_{-\infty}^{\infty}xf(x)dx + b = a\mathbb{E}[X]+b,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[g(X)+h(X)] &= \int_{-\infty}^{\infty}\left(g(x)+h(x)\right)f(x)dx \\ &= \int_{-\infty}^{\infty}g(x)f(x)dx + \int_{-\infty}^{\infty}h(x)f(x)dx\end{aligned}$$

Expectation and Variance

Solution (Properties of Expectation)

If X is continuous with pdf f, then

$$\begin{aligned}\mathbb{E}[aX+b] &= \int_{-\infty}^{\infty}(ax+b)f(x)dx \\ &= a\int_{-\infty}^{\infty} xf(x)dx + b = a\mathbb{E}[X]+b,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[g(X)+h(X)] &= \int_{-\infty}^{\infty}\left(g(x)+h(x)\right)f(x)dx \\ &= \int_{-\infty}^{\infty} g(x)f(x)dx + \int_{-\infty}^{\infty} h(x)f(x)dx \\ &= \mathbb{E}[g(X)]+\mathbb{E}[h(X)].\end{aligned}$$

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

$$\mathrm{Var}(aX+b) = \mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2]$$

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

$$\begin{aligned}\mathrm{Var}(aX+b) &= \mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2] \\ &= \mathbb{E}[(aX+b-(a\mu+b))^2]\end{aligned}$$

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

$$\begin{aligned}
\mathrm{Var}(aX+b) &= \mathbb{E}[(aX + b - \mathbb{E}[aX+b])^2] \\
&= \mathbb{E}[(aX + b - (a\mu + b))^2] \\
&= \mathbb{E}[a^2\,(X-\mu)^2]
\end{aligned}$$

Expectation and Variance

Solution (Properties of Variance)

Let $\mathbb{E}[X] = \mu$. Then

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

Also

$$\begin{aligned}
\mathrm{Var}(aX+b) &= \mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2] \\
&= \mathbb{E}[(aX+b-(a\mu+b))^2] \\
&= \mathbb{E}[a^2(X-\mu)^2] \\
&= a^2\mathrm{Var}(X).
\end{aligned}$$

Expectation and Variance

Definition (Moments)

The rth moment of a random variable X is $\mathbb{E}[X^r]$. Note that the expectation of X, or indeed any moment of X, need not be finite.

Question (Variance and Second Moments)

Prove that $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$. When do we get strict equality?

Solution

We have

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

But $\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \geq 0$, so $\mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$. Furthermore, the equality is strict if and only if $\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = 0$, which occurs when $X = \mathbb{E}[X]$, i.e., X is constant.

Important Continuous Distributions

pdf and cdf

Recall that a random variable X is said to have continuous distribution if there is a function f (called the pdf) so that

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx.$$

The cdf is then defined to be

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(u)du.$$

Important Continuous Distributions

Uniform Distribution

A random variable X is said to be *uniformly distributed* on the interval $[a, b]$ (denoted $X \sim \mathrm{U}[a, b]$) if the pdf is given by

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

Properties of the Uniform Distribution

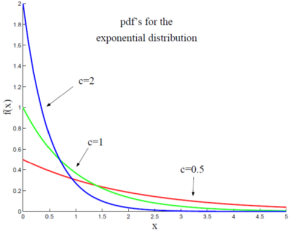
If $X \sim U[a, b]$ then

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

Important Continuous Distributions

Exponential Distribution

A random variable X is said to have an *exponential distribution* on $[0, \infty)$ with parameter λ (denoted $X \sim \text{Exp}(\lambda)$) if the pdf is given by $f(x) = \lambda e^{-\lambda x}$. The cdf is then $F(x) = 1 - e^{-\lambda x}$.



Properties of the Exponential Distribution

If $X \sim \text{Exp}(\lambda)$ then

- $\mathbb{E}[X] = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}$
- $\text{Var}[X] = \int_0^\infty x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

Important Continuous Distributions

The exponential distribution often arises for random variables X which describe 'how long do we have to wait until event Y happens?'. The exponential distribution is useful if the probability of the event Y happening in any time interval only depends on the length of the time interval, not the location of the time interval itself.

For example, 'how long until my favourite band releases their next song?' could be modeled with the exponential distribution.

The exponential distribution is the continuous analogue of 'if I keep tossing coins, how long until I get *Heads*?'

Important Continuous Distributions

Question (The Exponential Distribution is Memoryless)

If $X \sim \text{Exp}(\lambda)$, show that $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$ for any $s, t > 0$.

Solution

Important Continuous Distributions

Question (The Exponential Distribution is Memoryless)

If $X \sim \text{Exp}(\lambda)$, show that $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$ for any $s, t > 0$.

Solution

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

Important Continuous Distributions

Question (The Exponential Distribution is Memoryless)

If $X \sim \text{Exp}(\lambda)$, show that $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$ for any $s, t > 0$.

Solution

$$\begin{aligned}
\mathbb{P}(X > s + t | X > s) &= \frac{\mathbb{P}(X > s + t, X > s)}{\mathbb{P}(X > s)} \\
&= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)}
\end{aligned}$$

Important Continuous Distributions

Question (The Exponential Distribution is Memoryless)

If $X \sim \text{Exp}(\lambda)$, show that $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$ for any $s, t > 0$.

Solution

$$\begin{aligned}
\mathbb{P}(X > s + t | X > s) &= \frac{\mathbb{P}(X > s + t, X > s)}{\mathbb{P}(X > s)} \\
&= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}
\end{aligned}$$

Important Continuous Distributions

Question (The Exponential Distribution is Memoryless)

If $X \sim \text{Exp}(\lambda)$, show that $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$ for any $s, t > 0$.

Solution

$$\begin{aligned}
\mathbb{P}(X > s + t | X > s) &= \frac{\mathbb{P}(X > s + t, X > s)}{\mathbb{P}(X > s)} \\
&= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\
&= e^{-\lambda t} = \mathbb{P}(X > t)
\end{aligned}$$

Question (Probabilities of the Exponential Distribution)

Let $X \sim \text{Exp}(\lambda)$. For every $k > 0$ and non-negative integer k, find $\mathbb{P}(kd < X < (k+1)d)$.

Solution

Important Continuous Distributions

Question (Probabilities of the Exponential Distribution)

Let $X \sim \text{Exp}(\lambda)$. For every $k > 0$ and non-negative integer k, find $\mathbb{P}(kd < X < (k+1)d)$.

Solution

$$\mathbb{P}(kd < X < (k+1)d) = \mathbb{P}(X \leq (k+1)d) - \mathbb{P}(X \leq dk)$$

Important Continuous Distributions

Question (Probabilities of the Exponential Distribution)

Let $X \sim \text{Exp}(\lambda)$. For every $k > 0$ and non-negative integer k, find $\mathbb{P}(kd < X < (k+1)d)$.

Solution

$$\begin{aligned}\mathbb{P}(kd < X < (k+1)d) &= \mathbb{P}(X \leq (k+1)d) - \mathbb{P}(X \leq dk) \\ &= (1 - e^{-\lambda(k+1)d}) - (1 - e^{-\lambda kd})\end{aligned}$$

Important Continuous Distributions

Question (Probabilities of the Exponential Distribution)

Let $X \sim \mathrm{Exp}(\lambda)$. For every $k > 0$ and non-negative integer k, find $\mathbb{P}(kd < X < (k+1)d)$.

Solution

$$\begin{aligned}
\mathbb{P}(kd < X < (k+1)d) &= \mathbb{P}(X \leq (k+1)d) - \mathbb{P}(X \leq dk) \\
&= (1 - e^{-\lambda(k+1)d}) - (1 - e^{-\lambda kd}) \\
&= e^{-\lambda kd} - e^{-\lambda(k+1)d}
\end{aligned}$$

Question (Probabilities of the Exponential Distribution)

Let $X \sim \mathrm{Exp}(\lambda)$. For every $k > 0$ and non-negative integer k, find $\mathbb{P}(kd < X < (k+1)d)$.

Solution

$$\begin{aligned}
\mathbb{P}(kd < X < (k+1)d) &= \mathbb{P}(X \leq (k+1)d) - \mathbb{P}(X \leq dk) \\
&= (1 - e^{-\lambda(k+1)d}) - (1 - e^{-\lambda kd}) \\
&= e^{-\lambda kd} - e^{-\lambda(k+1)d} \\
&= e^{-\lambda kd}(1 - e^{-\lambda d})
\end{aligned}$$

Important Continuous Distributions

Definition (Gamma Function)

For $\alpha > 0$, the Gamma function is defined as
$\Gamma(\alpha) = \int_0^{\infty} u^{\alpha-1} e^{-u} du$.

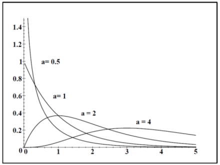
Properties of the Gamma Function

- $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$ for each $\alpha > 0$.
- $\Gamma(n) = (n - 1)!$ for each $n = 1, 2, 3, \cdots$.
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Important Continuous Distributions

Definition (Gamma Distribution)

We say that X has a *gamma distribution with shape parameter* α *and scale parameter* λ on the interval $[0, \infty)$ (denoted $X \sim \text{Gam}(\alpha, \lambda)$) if its pdf is given by $f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$.



Properties of the Gamma Distribution

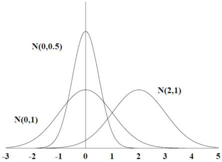
If $X \sim \text{Gam}(\alpha, \lambda)$ then

- $\mathbb{E}[X] = \frac{\alpha}{\lambda}$
- $\text{Var} = \frac{\alpha}{\lambda^2}$
- $X \sim \text{Exp}(\lambda)$ if $\alpha = 1$.

Important Continuous Distributions

Definition (Gaussian Distribution)

We say that X has a *Gaussian/normal distribution with parameters μ and σ^2* on $\mathbb{R}$ (denoted $X \sim \mathrm{N}(\mu, \sigma^2)$) if its pdf is given by $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$.



Properties of the Gaussian Distribution

If $X \sim \mathrm{N}(\mu, \sigma^2)$ then

- $\mathbb{E}[X] = \mu$
- $\mathrm{Var}[X] = \sigma^2$.