CS70: Jean Walrand: Lecture 35.

Continuous Probability 2

1. Review: CDF, PDF
2. Examples
3. Properties
4. Expectation
5. Expectation of Function
6. Variance
7. Independent Continuous RVs

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

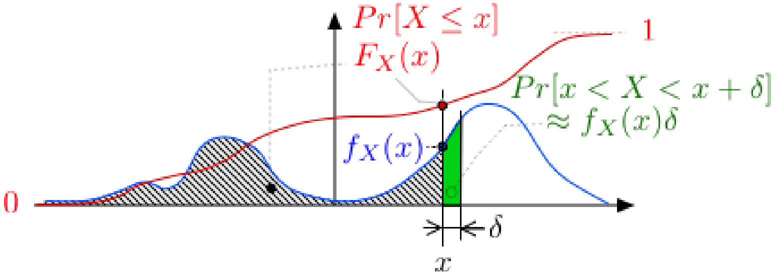
Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon \approx Pr[X \in (x, x+\varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function (CDF) of X.

$f_X(\cdot)$ is the probability density function (PDF) of X.

A Picture



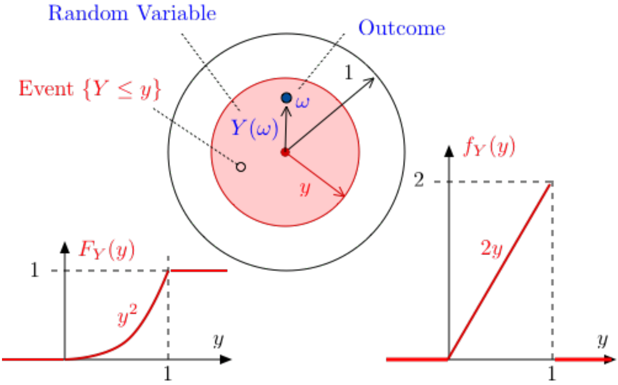
The pdf $f_X(x)$ is a nonnegative function that integrates to 1.

The cdf $F_X(x)$ is the integral of f_X.

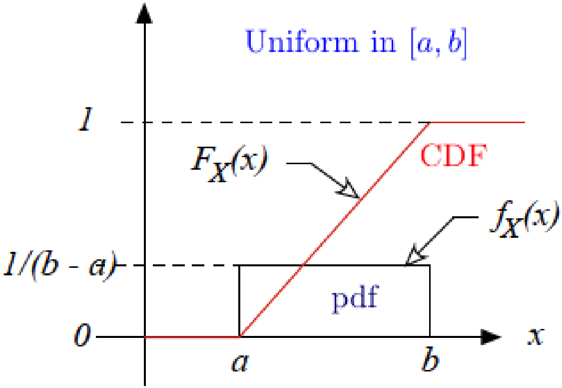
$$Pr[x < X < x + \delta] \approx f_X(x)\delta$$

$$Pr[X \leq x] = F_x(x) = \int_{-\infty}^{x} f_X(u)du$$

Target


Random Variable
Outcome
Event $\{Y \leq y\}$
1
ω
$Y(\omega)$
y
$f_Y(y)$
2
$2y$
$F_Y(y)$
1
y^2
y
1
y
1

$U[a, b]$

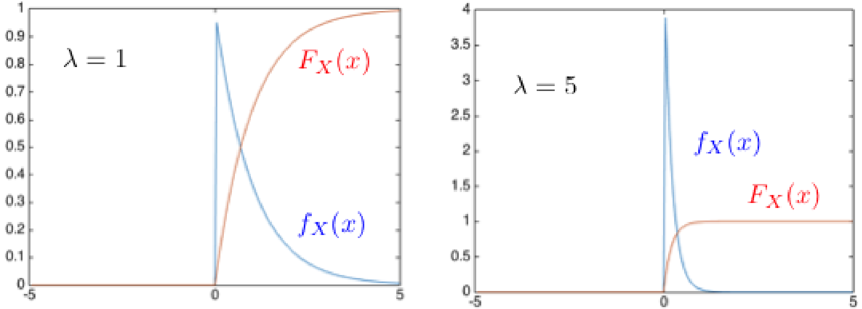


$Expo(\lambda)$

The exponential distribution with parameter $\lambda > 0$ is defined by

$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-\lambda x}, & \text{if } x \geq 0. \end{cases}$$



Note that $Pr[X > t] = e^{-\lambda t}$ for $t > 0$.

Some Properties

1. ***Expo*** **is memoryless.** Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

$$\begin{aligned} Pr[X > t + s \mid X > s] &= \frac{Pr[X > t + s]}{Pr[X > s]} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} \\ &= Pr[X > t]. \end{aligned}$$

'Used is a good as new.'

2. Scaling ***Expo*****.** Let $X = Expo(\lambda)$ and $Y = aX$ for some $a > 0$. Then

$$\begin{aligned} Pr[Y > t] &= Pr[aX > t] = Pr[X > t/a] \\ &= e^{-\lambda(t/a)} = e^{-(\lambda/a)t} = Pr[Z > t] \text{ for } Z = Expo(\lambda/a). \end{aligned}$$

Thus, $a \times Expo(\lambda) = Expo(\lambda/a)$.

Also, $Expo(\lambda) = \frac{1}{\lambda} Expo(1)$.

More Properties

3. Scaling Uniform. Let $X = U[0,1]$ and $Y = a + bX$ where $b > 0$. Then,

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = \frac{1}{b}\delta, \text{ for } 0 < \frac{y-a}{b} < 1 \\ &= \frac{1}{b}\delta, \text{ for } a < y < a+b. \end{aligned}$$

Thus, $f_Y(y) = \frac{1}{b}$ for $a < y < a+b$. Hence, $Y = U[a, a+b]$.

Replacing b by $b-a$ we see that, if $X = U[0,1]$, then $Y = a + (b-a)X$ is $U[a,b]$.

Some More Properties

4. Scaling pdf. Let $f_X(x)$ be the pdf of X and $Y = a + bX$ where $b > 0$. Then

$$
\begin{aligned}
Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\
&= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = f_X(\frac{y-a}{b})\frac{\delta}{b}.
\end{aligned}
$$

Now, the left-hand side is $f_Y(y)\delta$. Hence,

$$f_Y(y) = \frac{1}{b} f_X(\frac{y-a}{b}).$$

Expectation

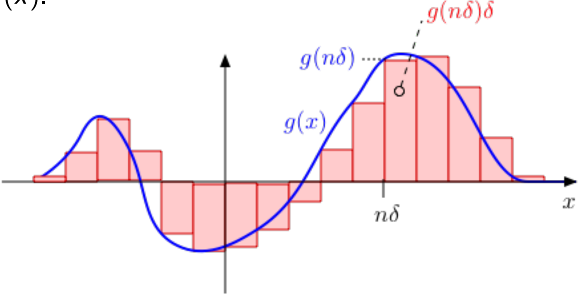
Definition: The **expectation** of a random variable X with pdf $f(x)$ is *defined* as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_n (n\delta) Pr[X = n\delta] = \sum_n (n\delta) f_X(n\delta)\delta = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Indeed, for any g, one has $\int g(x)dx \approx \sum_n g(n\delta)\delta$. Choose $g(x) = x f_X(x)$.



Examples of Expectation

1. $X = U[0,1]$. Then, $f_X(x) = 1\{0 \le x \le 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

2. $X =$ distance to 0 of dart shot uniformly in unit circle. Then $f_X(x) = 2x1\{0 \le x \le 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.2x dx = [\frac{2x^3}{3}]_0^1 = \frac{2}{3}.$$

Examples of Expectation

3. $X = Expo(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x \lambda e^{-\lambda x} dx = -\int_0^\infty x de^{-\lambda x}.$$

Recall the integration by parts formula:

$$\begin{aligned} \int_a^b u(x) dv(x) &= \left[u(x)v(x) \right]_a^b - \int_a^b v(x) du(x) \\ &= u(b)v(b) - u(a)v(a) - \int_a^b v(x) du(x). \end{aligned}$$

Thus,

$$\begin{aligned} \int_0^\infty x de^{-\lambda x} &= [xe^{-\lambda x}]_0^\infty - \int_0^\infty e^{-\lambda x} dx \\ &= 0 - 0 + \frac{1}{\lambda} \int_0^\infty de^{-\lambda x} = -\frac{1}{\lambda}. \end{aligned}$$

Hence, $E[X] = \frac{1}{\lambda}$.

Multiple Continuous Random Variables

One defines a pair (X, Y) of continuous RVs by specifying $f_{X,Y}(x,y)$ for $x, y \in \Re$ where

$$f_{X,Y}(x,y)dxdy = Pr[X \in (x, x+dx), Y \in (y+dy)].$$

The function $f_{X,Y}(x,y)$ is called the joint pdf of X and Y.

Example: Choose a point (X, Y) uniformly in the set $A \subset \Re^2$. Then

$$f_{X,Y}(x,y) = \frac{1}{|A|} 1\{(x,y) \in A\}$$

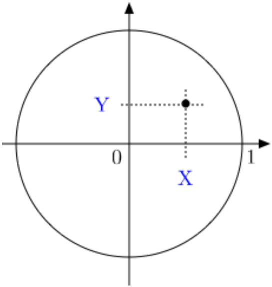
where $|A|$ is the area of A.

Interpretation. Think of (X, Y) as being discrete on a grid with mesh size ε and $Pr[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2$.

Extension: $\mathbf{X} = (X_1, \ldots, X_n)$ with $f_{\mathbf{X}}(\mathbf{x})$.

Example of Continuous (X, Y)

Pick a point (X, Y) uniformly in the unit circle.



Thus, $f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}$.

Consequently,

$$Pr[X > 0, Y > 0] = \frac{1}{4}$$
$$Pr[X < 0, Y > 0] = \frac{1}{4}$$
$$Pr[X^2 + Y^2 \leq r^2] = r^2$$
$$Pr[X > Y] = \frac{1}{2}.$$

Independent Continuous Random Variables

Definition: The continuous RVs X and Y are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A] Pr[Y \in B], \forall A, B.$$

Theorem: The continuous RVs X and Y are independent if and only if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

Proof: As in the discrete case.

Definition: The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1] \cdots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$

Theorem: The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if

$$f_{\mathbf{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

Proof: As in the discrete case.

Examples of Independent Continuous RVs

1. Minimum of Independent *Expo***.** Let $X = Expo(\lambda)$ and $Y = Expo(\mu)$ be independent RVs.

Recall that $Pr[X > u] = e^{-\lambda u}$. Then

$$\begin{aligned} Pr[\min\{X, Y\} > u] &= Pr[X > u, Y > u] = Pr[X > u]Pr[Y > u] \\ &= e^{-\lambda u} \times e^{-\mu u} = e^{-(\lambda+\mu)u}. \end{aligned}$$

This shows that $\min\{X, Y\} = Expo(\lambda + \mu)$.

Thus, the minimum of two independent exponentially distributed RVs is exponentially distributed.

2. Minimum of Independent $U[0,1]$**.** Let $X, Y = [0,1]$ be independent RVs. Let also $Z = \min\{X, Y\}$. What is f_Z?

One has

$$Pr[Z > u] = Pr[X > u]Pr[Y > u] = (1-u)^2.$$

Thus $F_Z(u) = Pr[Z \leq u] = 1 - (1-u)^2$.

Hence, $f_Z(u) = \frac{d}{du}F_Z(u) = 2(1-u), u \in [0,1]$. In particular, $E[Z] = \int_0^1 u f_Z(u)du = \int_0^1 2u(1-u)du = 2\frac{1}{2} - 2\frac{1}{3} = \frac{1}{3}$.

Expectation of Function of RVs

Definitions: (a) The expectation of a function of a random variable is *defined* as

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx.$$

(b) The expectation of a function of multiple random variables is defined as

$$E[h(\mathbf{X})] = \int \cdots \int h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n.$$

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$. Then,

$$E[h(X)] = \sum_n h(n\delta) Pr[X = n\delta] = \sum_n h(n\delta) f_X(n\delta)\delta = \int_{-\infty}^{\infty} h(x) f_X(x) dx.$$

Indeed, for any g, one has $\int g(x)dx \approx \sum_n g(n\delta)\delta$. Choose $g(x) = h(x)f_X(x)$.

The case of multiple RVs is similar.

Examples of Expectation of Function

Recall: $E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx$.

1. Let $X = U[0,1]$. Then
$$E[X^n] = = \int_0^1 x^n dx = [\frac{x^{n+1}}{n+1}]_0^1 = \frac{1}{n+1}.$$

2. Let $X = U[0,1]$ and $\theta > 0$. Then
$$E[\cos(\theta X)] = \int_0^1 \cos(\theta x) dx = [\frac{1}{\theta}\sin(\theta x)]_0^1 = \frac{\sin(\theta)}{\theta}.$$

3. Let $X = Expo(\lambda)$. Then
$$\begin{aligned} E[X^n] &= \int_0^\infty x^n \lambda e^{-\lambda x} dx = -\int_0^\infty x^n de^{-\lambda x} \\ &= -[x^n e^{-\lambda x}]_0^\infty + \int_0^\infty e^{-\lambda x} dx^n \\ &= \frac{n}{\lambda}\int_0^\infty x^{n-1}\lambda e^{-\lambda x} dx = \frac{n}{\lambda} E[X^{n-1}]. \end{aligned}$$
Since $E[X^0] = 1$, this implies by induction that $E[X^n] = \frac{n!}{\lambda^n}$.

Linearity of Expectation

Theorem Expectation is linear.

Proof: 'As in the discrete case.' □

Example 1: $X = U[a,b]$. Then

(a) $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}$. Thus,

$$E[X] = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} [\frac{x^2}{2}]_a^b = \frac{a+b}{2}.$$

(b) $X = a + (b-a)Y$, $Y = U[0,1]$. Hence,

$$E[X] = a + (b-a)E[Y] = a + \frac{b-a}{2} = \frac{a+b}{2}.$$

Example 2: X, Y are $U[0,1]$. Then

$$E[3X - 2Y + 5] = 3E[X] - 2E[Y] + 5 = 3\frac{1}{2} - 2\frac{1}{2} + 5 = 5.5.$$

Expectation of Product of Independent RVs

Theorem If X, Y, X are mutually independent, then

$$E[XYZ] = E[X]E[Y]E[Z].$$

Proof: Same as discrete case.

Example: Let X, Y, Z be mutually independent and $U[0,1]$. Then

$$\begin{aligned} E[(X+2Y+3Z)^2] &= E[X^2+4Y^2+9Z^2+4XY+6XZ+12YZ] \\ &= \frac{1}{3}+4\frac{1}{3}+9\frac{1}{3}+4\frac{1}{2}\frac{1}{2}+6\frac{1}{2}\frac{1}{2}+12\frac{1}{2}\frac{1}{2} \\ &= \frac{14}{3}+\frac{22}{4} \approx 10.17. \end{aligned}$$

Variance

Definition: The **variance** of a continuous random variable X is defined as

$$var[X] = E((X - E(X))^2) = E(X^2) - (E(X))^2.$$

Example 1: $X = U[0,1]$. Then

$$var[X] = E[X^2] - E[X]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Example 2: $X = Expo(\lambda)$. Then $E[X] = \lambda^{-1}$ and $E[X^2] = 2/(\lambda^2)$. Hence, $var[X] = 1/(\lambda^2)$.

Example 3: Let X, Y, Z be independent. Then

$$var[X + Y + Z] = var[X] + var[Y] + var[Z],$$

as in the discrete case.

Summary

Continuous Probability 2

1. pdf: $Pr[X \in (x, x+\delta]] = f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = \int_{-\infty}^{x} f_X(y)dy$.
3. $U[a,b]$, $Expo(\lambda)$, target.
4. Expectation: $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.
5. Expectation of function: $E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)dx$.
6. Variance: $var[X] = E[(X-E[X])^2] = E[X^2] - E[X]^2$.
7. $f_{\mathbf{X}}(\mathbf{x})dx_1 \cdots dx_n = Pr[X_1 \in (x_1, x_1 + dx_1), \ldots, X_n \in (x_n, x_n + dx_n)]$.
8. $X_1, \ldots, X_n$ are mutually independent iff $f_{\mathbf{X}} = f_{X_1} \times \cdots \times f_{X_n}$.
9. $\mathbf{X}$ mutually independent $\Rightarrow E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$.
10. $E[h(\mathbf{X})] = \int \cdots \int h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n$.
11. Expectation is linear.