Continuous random variables

Discrete random variable: takes values in a finite or countable set, e.g.
- $X \in \{1, 2, ..., 6\}$ with equal probability
- $X$ is a positive integer $i$ with probability $2^{-i}$

Continuous random variable: takes values in an uncountable set, e.g.
- $X$ is the weight of a random person (a real number)
- $X$ is a randomly selected point inside a unit square
- $X$ is the waiting time until the next packet arrives at the server

$f(x)$: $\mathbb{R} \rightarrow \mathbb{R}$, the probability density function (or simply “density”)

Require:
- $f(x) \geq 0$, and
- $\int_{-\infty}^{+\infty} f(x) \, dx = 1$
  - nonnegative, and
  - normalized,
  - just like discrete PMF

$F(x)$: the cumulative distribution function (aka the “distribution”)

$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x) \, dx$ (Area left of $a$)

$P(a < X \leq b) =$
F(x): the cumulative distribution function (aka the “distribution”)

\[ F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x) \, dx \]  

(Area left of a)

\[ P(a < X \leq b) = F(b) - F(a) \]  

(Area between a and b)

Relationship between f(x) and F(x)?

F(x): the cumulative distribution function (aka the “distribution”)

\[ F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x) \, dx \]  

(Area left of a)

\[ P(a < X \leq b) = F(b) - F(a) \]  

(Area between a and b)

A key relationship:

\[ f(x) = \frac{d}{dx} F(x), \text{ since } F(a) = \int_{-\infty}^{a} f(x) \, dx, \]

why is it called a density?

Densities are not probabilities; e.g. may be > 1

\[ P(X = a) = \lim_{\varepsilon \to 0} P(a-\varepsilon < X \leq a) = F(a)-F(a) = 0 \]

I.e., the probability that a continuous r.v. falls at a specified point is zero.

But

the probability that it falls near that point is proportional to the density:

\[ f(x) \]
Densities are not probabilities; e.g., may be > 1

\[ P(X = a) = \lim_{\varepsilon \to 0} P(a-\varepsilon < X \leq a) = F(a)-F(a) = 0 \]

I.e.,
   the probability that a continuous r.v. falls at a specified point is zero.

But
   the probability that it falls near that point is proportional to the density:

\[ P(a - \varepsilon/2 < X \leq a + \varepsilon/2) = \]
\[ F(a + \varepsilon/2) - F(a - \varepsilon/2) \approx \varepsilon \cdot f(a) \]

I.e., in a large random sample, expect more samples where density is higher (hence the name “density”).

Much of what we did with discrete r.v.s carries over almost unchanged, with \( \sum \) replaced by \( \int \)...

E.g.

For discrete r.v. \( X \), \[ E[X] = \sum x \cdot p(x) \]

For continuous r.v. \( X \), \[ E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx \]

Why?
      (a) We define it that way
      (b) The probability that \( X \) falls “near” \( x \), say within \( x \pm dx/2 \), is \( \approx f(x)dx \), so the “average” \( X \) should be \( \approx \sum xf(x)dx \) (summed over grid points spaced \( dx \) apart on the real line) and the limit of that as \( dx \to 0 \) is \( \int xf(x)dx \)

What is \( F(x) \)? What is \( E(X) \)?
Let \( f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases} \)

\[
F(a) = \int_{-\infty}^{a} f(x)dx = \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 < a \leq 1 \quad \text{(since } a = \int_{0}^{a} 1dx) \\ 1 & \text{if } 1 < a \end{cases}
\]

\[
E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{1} x dx = \frac{x^2}{2} \bigg|_{0}^{1} = \frac{1}{2}
\]

\[
E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_{0}^{1} x^2 dx = \frac{x^3}{3} \bigg|_{0}^{1} = \frac{1}{3}
\]

### Variance

Definition is same as in the discrete case:

\[
\text{Var}[X] = E[(X-\mu)^2] \quad \text{where } \mu = E[X]
\]

Identity still holds:

\[
\text{Var}[X] = E[X^2] - (E[X])^2
\]

### Linearity

- \(E[aX+b] = aE[X]+b\)
- \(E[X+Y] = E[X]+E[Y]\)

### Functions of a random variable

\[
E[g(X)] = \int g(x)f(x)dx
\]

Alternatively, let \(Y = g(X)\), find the density of \(Y\), say \(f_Y\), and directly compute \(E[Y] = \int y f_Y(y)dy\).

### Example

Let \( f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases} \)

\[
F(a) = \int_{-\infty}^{a} f(x)dx = \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 < a \leq 1 \quad \text{(since } a = \int_{0}^{a} 1dx) \\ 1 & \text{if } 1 < a \end{cases}
\]

\[
E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{1} x dx = \frac{x^2}{2} \bigg|_{0}^{1} = \frac{1}{2}
\]

\[
E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_{0}^{1} x^2 dx = \frac{x^3}{3} \bigg|_{0}^{1} = \frac{1}{3}
\]

Identity still holds:

\[
\text{Var}[X] = E[(X-\mu)^2] = E[X^2] - (E[X])^2
\]
Let \( f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases} \)

\[
F(a) = \int_{-\infty}^{a} f(x) \, dx = \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 < a \leq 1 \text{ (since } a = \int_{0}^{a} 1 \, dx) \\ 1 & \text{if } 1 < a \end{cases}
\]

\[
E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{0}^{1} x \, dx = \frac{x^2}{2} \bigg|_{0}^{1} = \frac{1}{2}
\]

\[
E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) \, dx = \int_{0}^{1} x^2 \, dx = \frac{x^3}{3} \bigg|_{0}^{1} = \frac{1}{3}
\]

\[
\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \quad (\sigma \approx 0.29)
\]

Continuous random variable \( X \) has density \( f(x) \), and

\[
\Pr(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx
\]

\[
E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx
\]

\[
E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x) \, dx
\]

**uniform random variables**

\( X \sim \text{Uni}(\alpha, \beta) \) is uniform in \([\alpha, \beta]\)

\[
f(x) = \begin{cases} \frac{1}{\beta - \alpha} & x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}
\]

\[
\Pr(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx = \frac{b - a}{\beta - \alpha}
\]

\[
E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx = \frac{\alpha + \beta}{2}
\]
waiting for “events”

Radioactive decay: How long until the next alpha particle?

Customers: how long until the next customer/packet arrives at the checkout stand/server?

Buses: How long until the next #71 bus arrives on the Ave?

Yes, they have a schedule, but given the vagaries of traffic, riders with bikes and baby-carriages, etc., can they stick to it?

Assuming events are independent, happening at some fixed average rate of λ per unit time – the waiting time until the next event is exponentially distributed (next slide)

exponential random variables

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

The Exponential Density Function

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

Memorylessness:

$$\Pr(X \geq t) = e^{-\lambda t} = 1-F(t)$$

$$\Pr(X > s + t | X > s) = \Pr(X > t)$$

Assuming exp distr, if you’ve waited $s$ minutes, prob of waiting $t$ more is exactly same as $s = 0$

Relation to Poisson

Same process, different measures:

Poisson: how many events in a fixed time;

Exponential: how long until the next event

$\lambda$ is avg # per unit time;

$1/\lambda$ is mean wait
Time it takes to check someone out at a grocery store is exponential with an average value of 10. If someone arrives to the line immediately before you, what is the probability that you will have to wait between 10 and 20 minutes?

\[ T \sim \exp(10^{-1}) \]

\[ Pr(10 \leq T \leq 20) = \int_{10}^{20} \frac{1}{10} e^{-\frac{x}{10}} \, dx \]

\[ y = \frac{x}{10} \quad dy = \frac{1}{10} \, dx \]

\[ Pr(10 \leq T \leq 20) = \int_{1}^{2} e^{-y} \, dy = -e^{-y}\bigg|_{1}^{2} = (e^{-1} - e^{-2}) \]

Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with an average value of 10,000 miles. If the car has already been used for 2000 miles, and the owner wants to take a 5000 mile trip, what is the probability she will be able to complete the trip without replacing the battery?

\[ N \sim \exp(1/10,000) \]

\[ Pr(N \geq 7000|N \geq 2000) = \frac{Pr(N \geq 7000)}{Pr(N \geq 2000)} \]

\[ Pr(N \geq 7000) = e^{-7000/10000} \]

\[ Pr(N \geq 2000) = e^{-2000/10000} \]

answer \( = e^{-5000/10000} = e^{-0.5} \)

---

**normal random variables**

\( X \) is a normal (aka Gaussian) random variable  \( X \sim N(\mu, \sigma^2) \)

\[ f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

\[ E[X] = \mu \quad \text{Var}[X] = \sigma^2 \]

The Standard Normal Density Function

changing \( \mu, \sigma \)

The graph shows how the normal density function changes with different \( \mu \) and \( \sigma \) values:

- When \( \mu = 0 \) and \( \sigma = 1 \), the curve is centered at 0 and has a standard deviation of 1.
- When \( \mu = 4 \) and \( \sigma = 2 \), the curve is shifted to the right by 4 units and has a standard deviation of 2.

The density at \( \mu \) is \( \approx 0.399/\sigma \).
X is a normal random variable \( X \sim N(\mu, \sigma^2) \)

\[ Y = aX + b \]

\[ E[Y] = E[aX+b] = a\mu + b \]

\[ \text{Var}[Y] = \text{Var}[aX+b] = a^2\sigma^2 \]

\[ Y \sim N(a\mu + b, a^2\sigma^2) \]

Important special case: \( Z = (X-\mu)/\sigma \sim N(0,1) \)

\[ Z \sim N(0,1) \] "standard (or unit) normal"

Use \( \Phi(z) \) to denote CDF, i.e.

\[ \Phi(z) = \Pr(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx \]

no closed form 😞
X normal with mean 3 and variance 9.

What is

\[ Pr(X > 0) \]

\[ Pr(2 < X < 5) \]

\[ Pr(|X - 3| > 6) \]

X normal with mean 5 and variance \( \sigma^2 \)

If \( Pr(X > 9) = 0.2 \), then approximately what is \( \sigma^2 \)?

\[
Pr(X > 9) = Pr\left(\frac{X - 5}{\sigma} > \frac{9 - 5}{\sigma}\right) = 0.2
\]

\[
1 - \Phi\left(\frac{9 - 5}{\sigma}\right) = 0.2
\]

\[
\Phi\left(\frac{9 - 5}{\sigma}\right) = 0.8
\]

Look up in N(0,1) table to find our \( \nu \) gives \( \Phi(\nu) = 0.8 \)

Set \( \frac{9 - 5}{\sigma} = \nu \) and solve for \( \sigma \)
pdf vs cdf

\[ f(x) = \frac{d}{dx} F(x) \quad F(a) = \int_{\infty}^{a} f(x) \, dx \]

sums become integrals, e.g.

\[ E[X] = \sum x \cdot p(x) \quad E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx \]

most familiar properties still hold, e.g.

\[ E[aX+bY+c] = aE[X]+bE[Y]+c \]

\[ \text{Var}[X] = E[X^2] - (E[X])^2 \]

Three important examples

\( X \sim \text{Uni}(\alpha, \beta) \) uniform in \([\alpha, \beta]\)

\[
\begin{align*}
f(x) &= \begin{cases} 
\frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\
0 & \text{otherwise}
\end{cases} \\
E[X] &= (\alpha+\beta)/2 \\
\text{Var}[X] &= (\alpha-\beta)^2/12
\end{align*}
\]

\( X \sim \text{Exp}(\lambda) \) exponential

\[
\begin{align*}
f(x) &= \begin{cases} 
\lambda e^{-\lambda x} & x \geq 0 \\
0 & x < 0
\end{cases} \\
E[X] &= \frac{1}{\lambda} \\
\text{Var}[X] &= \frac{1}{\lambda^2}
\end{align*}
\]

\( X \sim \text{N}(\mu, \sigma^2) \) normal (aka Gaussian, aka the big Kahuna)

\[
\begin{align*}
f(x) &= \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \\
E[X] &= \mu \\
\text{Var}[X] &= \sigma^2
\end{align*}
\]