Introduction to probability

Amit Tripathi

Indian Institute of Technology, Hyderabad

August 22, 2019

Continuous random variables

Definition

A random variable X is said to be a **continuous random variable** if there exists a function f satisfying the following

1. $f(x) \geq 0$ for all x.
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.
3. For any $a \leq b$, we have

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Definition

A function f as above is called the probability density function (PDF) of X.

Probability density function

Question

For a continuous random variable X, is cumulative distribution function continuous ? Is it differentiable ?

Relation between cumulative distribution function and probability density function

By definition, cumulative distribution function

$$F(a) := P(X \leq a)$$

Thus if X has pdf $f(x)$ then

$$F(a) := P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

Differentiating both sides w.r.t a yields

$$\frac{d}{da}F(a) = f(a)$$

Relation between cumulative distribution function and probability density function

Example

Let X be a continuous RV with DF F_X and PDF (probability density function) f_X. Find the PDF of $Y = X^2$.
By definition

$$F_Y(a) = P(X^2 \leq a) = \begin{cases} F_X(\sqrt{a}) - F_X(-\sqrt{a}) & a > 0 \\ 0 & a \leq 0 \end{cases}$$

Differentiating both sides w.r.t a we get

$$f_Y(a) = \begin{cases} \dfrac{f_X(\sqrt{a}) + f_X(-\sqrt{a})}{2\sqrt{a}} & a > 0 \\ 0 & a \leq 0 \end{cases}$$

Expectation of a continuous random variable

Recall that we defined expectation of discrete random variable as

$$E[X] = \sum_x xP(X = x)$$

For continuous case, there is no probability mass function but over a small interval Δx we can approximately say that (if $f(x)$ is pdf) then

$$f(x)d\,x \simeq P(x \leq X \leq x + \Delta x)$$

Thus analogously, we define

Definition

$E[X] = \int_{-\infty}^{\infty} xf(x)dx$ when X is a continuous random variable.

Expectation of a continuous random variable

Example

Suppose the density function of X is

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $E[X]$?

Expectation of a continuous random variable

Example

Suppose the density function of X is

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $E[\exp(X)]$? We proceed in the following manner:

1. We will first compute the probability distribution function of $Y = \exp X$.
2. From DF we will get density function of Y (taking derivative).
3. Finally the expectation of Y using the density function.

Expectation of a continuous random variable

Alternatively, we can use

Theorem

Let X be a continuous random variable with PDF $f(x)$. Assume that for some function g, $g(X)$ also defines a random variable. Then

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

Expectation of a continuous random variable

We revisit the previous example again:

Example

Suppose the density function of X is

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $E[\exp(X)]$?

Computing this time using the Theorem above,

$$E[\exp(X)] = \int_{-\infty}^{\infty} \exp(X) f(x)\, dx = \int_0^1 \exp(X)\, dx = e - 1$$

Variance of a continuous random variable

As before, **variance of** X is defined as

$$Var(X) = E[(X - \mu)^2]$$

and it can be calculated as

$$Var(X) = E[X^2] - \mu^2$$

Uniform random variables

Definition

If the probability density function $f(x)$ of a random variable is a nonzero constant λ over an interval (a, b) (with $a \neq b$) and zero elsewhere, we say that it is a **uniform random variable**.

Uniform random variables

Since $1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_a^b \lambda\,dx = \lambda(b-a)$. Thus

$$\lambda = \frac{1}{b-a}$$

Therefore we have

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

Uniform random variables

The cumulative distribution function is given as

$$F(x) = \int_{-\infty}^{x} f(y)\, dy = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$$

Uniform random variables

Example

Suppose X is uniformly distributed over $[a, b]$ and let $[c, d] \subset [a, b]$. Find the probability that $P(c \leq X \leq d)$?

Answer : $\dfrac{d - c}{b - a}$.

Expectation of uniform random variables

Question: If X is uniformly distributed over (a, b). What do we "expect" the expected value of X to be?

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_a^b \frac{x}{b-a}\, dx = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

Variance of uniform random variables

$$Var(X) = E[X^2] - E[X]^2 = \int_a^b \frac{x^2}{b-a}\,dx - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

Uniform random variable

Example

Consider a stick of unit length which is randomly broken at a point u whose distribution is uniform on the interval $(0, 1)$. Fix a point $p \in [0, 1]$. Find the expected length of the part of the broken stick which contains the point p.

We define a random variable X_p which denotes the length of the part of the broken stick which contains the point p when the stick is broken at point u. Then

$$X_p(u) = \begin{cases} u, & p < u \\ 1 - u, & p \geq u \end{cases}$$

Uniform random variable

Example

Now to compute $E[X_p]$ we identify $X_p(u)$ as a function of the random variable u with uniform distribution. Then

$$E[X_p] = \int_0^1 X_p(u) \cdot 1 \, d\,u = \int_0^p (1-u)\, d\,u + \int_p^1 u\, d\,u$$

$$= p - \frac{p^2}{2} + \frac{u^2}{2}\Big|_p^1 = p - \frac{p^2}{2} + \frac{1-p^2}{2} = \frac{1}{2} + p(1-p)$$

Exponential random variables

Definition

A continuous random variable whose probability density function is given as

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

is known as an **exponential random variable** with parameter λ.

Question

Can λ be negative?

Exponential random variables

$f(x)$ indeed defines a density function as

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^{\infty} \lambda\, e^{-\lambda\, x}\, dx = \int_0^{\infty} e^{-t}\, dt = 1$$

<u>**Cumulative distribution function**</u> is

$$F(a) = P(X \leq a) = \begin{cases} \int_0^a \lambda\, e^{-\lambda\, x}\, dx = 1 - e^{-\lambda a} & \text{when } a \geq 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

Exponential - expectation and variance

Lemma

If X is an exponential random variable with parameter λ then

1. $E[X^n] = \dfrac{n!}{\lambda^n}$.
2. $E[X] = \dfrac{1}{\lambda}$.
3. $Var(X) = \dfrac{1}{\lambda^2}$.

Exponential random variables

We start with

$$
\begin{aligned}
E[X^n] &= \int_0^\infty x^n \lambda \exp(-\lambda x) dx \\
&= \left(x^n \int \lambda \exp(-\lambda x) dx \right) \Big|_0^\infty - \int_0^\infty \left(\int \lambda \exp(-\lambda x) dx \right) \frac{dx^n}{dx} dx \\
&= 0 + \frac{n}{\lambda} E[X^{n-1}] \\
&= \frac{n}{\lambda} E[X^{n-1}]
\end{aligned}
$$

Thus we get

$$E[X^n] = \frac{n}{\lambda} E[X^{n-1}]$$

Putting $n = 1$ gives

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

Putting $n = 2$ gives

$$E[X^2] = \frac{2}{\lambda^2}$$

Therefore

$$Var(X) = E[X^2] - E[X]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Memoryless distributions

Definition

We say a random variable X is **memoryless** if

$$P(X > s + t \,|\, X > t) = P(X > s), \qquad \text{for all } s, t \geq 0$$

Essentially this means that the waiting time doesn't matter. Examples - time to next high magnitude earthquake/extreme weather phenomenon/accident on a road etc.

Memoryless distributions

Question

Is an exponential distribution memoryless?

Exponential random variables

Lemma

Suppose X is a memoryless random variable then X has an exponential distribution.

Being memoryless translates to

$$P(X > s + t) = P(X > s)P(X > t)$$

If we define $G(s) = 1 - F(s)$ for all s then this can be rewritten as

$$G(s + t) = G(s)G(t)$$

Exponential random variables

Proof.

We get (using continuity property)

$$G(s) = G(1)^s, \quad \text{for all } s \in \mathbb{R}$$

Since $G(1) = G(1/2)^2 \geq 0$, we can calculate the logarithm of $G(1)$. Putting $\lambda = -Log(G(1))$ we get

$$G(x) = \exp(-\lambda x)$$

In particular, the cumulative distribution function of X is given as

$$F(X \leq x) = 1 - G(x) = 1 - \exp(-\lambda x)$$

$\square$

Exponential distribution example

Example

Suppose that the time a electronic component can run before breaking down is exponentially distributed with an average value of N days. What is the probability that it will not break down in next n days.

Question: Do we need to know for how long the component has been working?

Exponential distribution example

Example

Question: What is the value of λ ?

Let X denotes the time that the component works before breaking down.

$$P(X < a) = 1 - \exp(-\lambda a)$$

Thus the probability that the component will not break down in n days is

$$P(X > n) = \exp\left(-\frac{n}{N}\right)$$

Exponential distribution

Suppose $E_1, E_2 \cdots$, denote a collection of "rare events" (governed by Poisson distribution). Consider a collection of random variables $\{T_0, T_1, \cdots, \}$ defined as follows:

$$T_i = \text{Time } t \text{ at which } i\text{'th event arrive}$$

Note here that each T_i is indexed by a discrete set but is in itself a continuous random variable.

Exponential distribution

We now find the distribution of inter-arrival times $T_i - T_{i-1}$ for $i = 1, 2, \cdots$ i.e. the distribution of time between two successive events satisfying the above hypothesis. Define

$$M_i = T_i - T_{i-1}$$

With this notation, we have

Theorem

Let N be a Poisson process with rate λ. Then the inter-arrival time M_i are independent random variables each having an exponential distribution with parameter λ.

Exponential distribution example

Example

Recall that if we have N_t number of some radioactive atoms at time t then the decay is predicted by equation

$$\frac{d\,N_t}{d\,t} = -\lambda N_t \tag{1}$$

Assuming that at time $t = 0$ there were N_0 atoms, we can solve this differential equation as

$$N_t = N_0 e^{-\lambda t}$$

This is the macroscopic approach which assumes that equation (1) is the governing law for the decay.

Exponential distribution example

Example

We now follow a microscopic approach at individual atom level. Instead of equation (1) - we now assume two elementary facts which have basis in quantum mechanics.

1. The event that atom i has decayed is independent of the event that atom j has decayed.
2. A radioactive atom has no knowledge of its history about when to decay. It is memoryless i.e. if we denote X_i as the time at which the i'th atom decays then

$$P(X_i > s + t \,|\, X_i > s) = P(X_i > t)$$

Exponential distribution example

Example

The only distribution which is memoryless is exponential distribution. Thus for a fixed atom, we get

$$P(X > t) = e^{-\lambda t}$$

At the beginning there were a total of N_0 atoms all in undecayed state. After time t, every atom has two possibility - either it has decayed (with probability $1 - e^{-\lambda t}$) or it is still undecayed (with probability $e^{-\lambda t}$).

Exponential distribution example

Example

If S denotes the total number of undecayed atoms then (using the elementary fact (1) above)

$$P(S = N) = \frac{N_0!}{N!(N_0 - N)!}(1 - e^{-\lambda t})^{N_0 - N} e^{-N\lambda t}$$

This is a Binomial distribution with parameters $(N_0, e^{-\lambda t})$. Thus the expected value of S is given as

$$E[S] = N_0 e^{-\lambda t}$$

which is nothing but the number of atoms predicted using macroscopic approach earlier.

Exponential distribution example

Some examples of physical phenomenon which satisfy these hypothesis:

1. Number of earthquakes in a given time interval.
2. Number of births/deaths in a small town (with reasonable population).
3. Radioactive decay etc.

Question

Question

Suppose X takes only non-negative integral values and is memoryless i.e. $P(X > m + n \,|, X > m) = P(X > n)$ for $m, n \geq 0$. Find the distribution of X.

Question

If exponential distribution governs the inter-arrival time for Poissonian events, then which distribution governs the inter-arrival time for Binomial events?

Joint distribution

There are scenarios where one is interested in more than one random variable. E.g. a weather scientist might be interested to know distribution of temperature as well as pressure together.

Definition

Let X and Y be two random variables. We define **joint cumulative distribution function** of X and Y to be

$$F(a, b) = P(X \leq a, Y \leq b) \qquad -\infty < a, b < \infty.$$

Joint distribution

Question: How to obtain the (marginal) distribution of X (or Y) from the joint distribution of X and Y?

We use the equality of events

$$\{X \leq a\} = \{X \leq a, Y < \infty\} = \bigcup_{b=1}^{\infty} \{X \leq a, Y \leq b\}$$

Therefore

$$\begin{aligned} F_X(a) = P(X \leq a) &= P(\bigcup_{b=1}^{\infty} \{X \leq a, Y \leq b\}) \\ &= \lim_{b \to \infty} P(X \leq a, Y \leq b) \\ &= \lim_{b \to \infty} F(a, b) \end{aligned}$$

Question: Why does the above limit exists?

Joint distribution

Definition

Let X and Y be two discrete random variable. We define **joint probability mass function** of X and Y to be

$$f(m,n) = P(X = m, Y = n)$$

Question: What is $\sum_m \sum_n P(X = m, Y = n)$?

Question: What is $\sum_n P(X = m, Y = n)$?

Joint distribution

Therefore, it is possible to recover probability mass function of X (or Y) from joint probability mass function as

$$P(X = m) = \sum_{n:p(m,n)>0} P(X = m, Y = n)$$

Joint distribution

Example

Consider an experiment where a fair dice is thrown and a fair coin is tossed.

1. Let X be the random variable equaling the output of dice and
2. Let Y be the random variable associated to output of the coin, defined as $Y(H) = 1$ and $Y(T) = 0$.

Then

$$P(X = 3, Y = 0) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

Joint distribution function can be calculated as

$$P(X \leq 3, Y \leq 0) = \sum_{i \leq 3} P(X = i, Y = 0) = \frac{3}{12} = \frac{1}{4}$$

Joint distribution

Example

Suppose that a balls are withdrawn (randomly and without replacement) from an urn containing r red, w white, and b blue balls.

Question: Let X and Y denote, respectively, the number of red and white balls chosen, then what is the joint probability mass function $P(X = i, Y = j)$?

$$P(X = i, Y = j) = \frac{\binom{r}{i}\binom{w}{j}\binom{b}{a-i-j}}{\binom{r+w+b}{a}}$$

Joint distribution

Definition

We will say that X and Y are **jointly continuous** if there exists a function f defined for all real x and y, such that for every set $C \subset \mathbb{R}^2$ we have

$$P(\{X, Y\} \in C) = \iint_{(x,y)\in C} f(x, y)\, d\, x\, d\, y$$

Definition

Such a function $f(x, y)$ is called the **joint density function** of X and Y.

Joint distribution

In particular for any a, b,

$$P(\{X \in (-\infty, a], Y \in (-\infty, b]\}) = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x, y) d\,x\,d\,y$$

Taking partial derivative with respect to a and b we get joint density function as

$$f(a, b) = \frac{\partial^2 F(a, b)}{\partial a\, \partial b}$$

Joint distribution

Question: How to obtain the (marginal) distribution of X (or Y) from the joint distribution of X and Y?

Again, we use the equality of events

$$\{X \leq a\} = \{X \leq a, Y < \infty\}$$

$$\int_a^\infty f_X(x)\, dx = \int_a^\infty \int_{-\infty}^\infty f(x, y) dx\, dy$$

Since this holds for all a, on comparing the integrands we get

$$f_X(x) = \int_{-\infty}^\infty f(x, y) d\, y$$

Joint distribution

To sum up: We get distribution of X from the joint distribution of X and Y by "integrating Y out" i.e.

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) d y$$

Joint distribution

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Let $a, b \geq 0$. Compute $P(X > a, Y < b)$.

By definition $P(X > a, Y < b)$ means $X \in (a, \infty)$ and $Y \in (0, b)$. Thus we get

$$P(X > a, Y < b) = \int_a^\infty \int_0^b 2e^{-x}e^{-2y}\,dx\,dy = e^{-a}(1 - e^{-2b})$$

Example

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Compute $P(X < Y)$.

By definition

$$\begin{aligned} P(X < Y) &= \int_0^\infty \int_0^y 2e^{-x}e^{-2y}\,dxdy \\ &= \int_0^\infty 2(1 - e^{-y})e^{-2y}\,dy \\ &= \int_0^\infty 2e^{-2y}\,dy - \int_0^\infty 2e^{-3y}\,dy = 1 - 2/3 = 1/3 \end{aligned}$$

Joint distribution

Example

The joint density function of X and Y are given as

$$f_{XY}(x,y) = \begin{cases} \frac{1}{2}\exp(-\frac{x+y}{2}), & 0 \le x \le y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Find the density function of X ? We "integrate Y out"

$$f_X(x) = \int_x^\infty f(x,y)dy = \exp(-x)$$

Joint distribution

Example

Suppose the joint density function of X and Y is

$$f(x,y) = \begin{cases} \lambda xy, & 0 < x < 1, 0 < y < 1, 0 < x + y < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find λ.

Uniform random variable - 2 dimensional version

An important application of uniform distribution is to geometry.

Definition

The random vector (X, Y) is said to be uniformly distributed over a region A in the plane if, for some constant c, its joint density is

$$f(x,y) = \begin{cases} c, & (X,Y) \in A \\ 0, & \text{otherwise} \end{cases}$$

Question

What is the value of c ?

Uniform random variable - 2 dimensional version

Example

Suppose we choose a point uniformly over a circle of radius R. The joint density function is given as

$$f(x, y) = \begin{cases} \dfrac{1}{\pi R^2}, & x^2 + y^2 \leq R^2 \\ 0, & \text{otherwise} \end{cases}$$

Uniform random variable - 2 dimensional version

Example

Question 1: What is the marginal density function of X? We "integrate out Y" to find it.

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} f(x,y)dy = \frac{2\sqrt{R^2 - x^2}}{\pi R^2}$$

(when $x^2 \leq R^2$).

Uniform random variable - 2 dimensional version

Example

Question 2: What is the probability that distance D of the point selected (from the origin) is less than a where $a < R$? Suppose A be the concentric circle of radius a, then this is given as

$$P(D \leq a) = \int_{(x,y)\in A} f(x,y)dx\,dy = \frac{A}{\pi R^2} = \frac{a^2}{R^2}$$

Uniform random variable - 2 dimensional version

Example

Question 3: Find $E[D]$? Density function of D is obtained by differentiating CDF w.r.t a

$$f_D(a) = \frac{d\,F(a)}{d\,a} = \frac{2a}{R^2}$$

Thus we get

$$E[D] = \int_0^R a \cdot \frac{2a}{R^2} da = \frac{2R}{3}$$

Uniform distribution

Example

Two numbers are chosen randomly (uniformly) from the interval $(0, 1)$. Find the probability that the distance between them is less than $\frac{1}{2}$.

1. What is the joint probability density function?
2. What is the area of interest?

Independent random variables

Suppose X, Y be two random variables. Let $F(a, b)$ denote the joint DF of X and Y. Let F_X (resp. F_Y) denote the DF of X (resp. Y).

Definition

We say X and Y are **independent** if for any sets $A, B \subset \mathbb{R}$

$$P(X \in A, Y \in B) = P(X \in A)P(X \in B)$$

Independent random variables

This in turn is equivalent to saying that for any $a, b \in \mathbb{R}$

$$F_{X,Y}(a,b) = F_X(a)F_Y(b)$$

Equivalently

$$\int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x,y)dx\,dy = \int_{-\infty}^{a} f_X(x)dx \int_{-\infty}^{b} f_Y(y)dy$$

i.e. (on comparing the integrands)

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

Expected value of product of independent random variables

Lemma

Let X and Y be independent random variables then

$$E[XY] = E[X]\,E[Y].$$

Expected value of areas

Example

One student selects a number U uniformly in the interval $(0, 1)$ and draws a square with sides U.

Another student selects two numbers V, W uniformly (and independently of each other) in the same interval $(0, 1)$ and draws a rectangle with sides V and W.

On an average, which figure will have the larger area - the square or the rectangle?

Revision slide

Question: Suppose the cumulative distribution function is given as

$$F_{X,Y}(x,y) = \begin{cases} (1-\exp(-x))(1-\exp(-2y)), & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

1. Find the cumulative distribution of X.
2. Find the joint density function of X and Y.
3. Are X and Y independent.

Normal random variable

Definition

We say that X is a normal random variable (or simply X is normally distributed) with parameters μ and σ^2 if the density function of X is given as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Observe that this function is symmetric about μ.

Notation

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

Normal random variable

Definition

We say that X is a **standard normal random variable** if $\mu = 0$ and $\sigma^2 = 1$. In this case, the density function of X is given as

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp^{-\frac{x^2}{2}}$$

Normal random variable

We will show that $f(x)$ is indeed a probability density function i.e., we must show that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\dfrac{x^2}{2}} dx = 1$$

Let $I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\dfrac{x^2}{2}} dx$.

Normal random variable

Then

$$I^2 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\frac{y^2}{2}} dy$$

which can be rewritten as

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp^{-\frac{x^2+y^2}{2}} dx\, dy$$

Normal random variable

We substitute $x = r\cos\theta$, $y = r\sin\theta$. Then

$$J = \det \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\[2ex] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = r$$

Normal random variable

Therefore $dy\, dz = J\, dr\, d\theta = r\, dr\, d\theta$ and hence

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp^{-\frac{x^2+y^2}{2}} \, dx\, dy$$

$$= \frac{1}{2\pi} \int_{0}^{\infty} \int_{0}^{2\pi} r \, \exp^{-\frac{r^2}{2}} \, dr\, d\theta$$

$$= \frac{1}{2\pi} \cdot 2\pi = 1$$

Therefore $I = 1$ as claimed.

De Moivre-Laplace's theorem

Theorem

Let X_n denotes a Binomial distribution with parameters (n, p). Suppose $a < b$ be two numbers. Then

$$\lim_{n \to \infty} P\left(a < \frac{X_n - np}{\sqrt{npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp^{-x^2/2} dx$$

Standard normal random variable - expectation

We compute mean for standard normal random variable:

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x\, e^{-\dfrac{x^2}{2}}\, dx$$

As the integrand is an odd function, we get

$$E[X] = 0$$

Standard normal random variable - variance

We now compute the variance. Since $E[X] = 0$, therefore $Var(X) = E[X^2]$.

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \exp^{-\frac{x^2}{2}}\, dx$$

Therefore

$$E[X^2] = \frac{1}{\sqrt{2\pi}} \left(-x \exp^{-\frac{x^2}{2}} |_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \exp^{-\frac{x^2}{2}}\, dx \right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx = 1$$

Standard normal random variable

So if X is a standard normal random variable, i.e. the density function of X is given as

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp^{-\frac{x^2}{2}}$$

Then,

$$E[X] = 0, \quad \text{and} \quad Var(X) = 1$$

Normal random variable

Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$. Define $Y = \dfrac{(X-\mu)}{\sigma}$.

Question: What is the density function of Y ?

$$P(Y \leq y) = P(\frac{X-\mu}{\sigma} \leq y) = P(X \leq \sigma y + \mu)$$

Since $X \sim \mathcal{N}(\mu, \sigma^2)$ we get

$$P(Y \leq y) = \int_{-\infty}^{\sigma y+\mu} \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\dfrac{(x-\mu)^2}{2\sigma^2}}\, dx$$

Normal random variable

Making the obvious substitution $y = \dfrac{(x-\mu)}{\sigma}$, we get

$$P(Y \leq y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp^{-\dfrac{y^2}{2}} \, dy$$

Question: What is $E[Y]$? Therefore what is $E[X]$ when $Y = \dfrac{(X-\mu)}{\sigma}$?

Question: What is $Var[Y]$? Therefore what is $Var[X]$?

Sum of random variables

Example

Suppose X and Y are discrete random variables with probability mass function $P(X = i) = P(Y = i) = 1/N$ for $i = 1, \cdots, N$ and 0 otherwise.

Question: Find probability mass function of $X + Y$?

E.g. we have

$$P(X + Y = 2) = P(X = 1, Y = 1)$$

Sum of random variables

Example

Assuming independence,

$$P(X+Y=3) = P(X=2, Y=1) + P(X=1, Y=2) = 2/N^2$$

and similarly

$$P(X+Y=4) = 3/N^2$$

Sum of random variables

Suppose X and Y are two **independent** and continuous random variables with probability density function f_X and f_Y respectively.

Question: What is the density function of $X + Y$?

$$F_{X+Y}(a) = P(X + Y \leq a) = \iint\limits_{x+y\leq a} f_{X,Y}(x, y)\, dx\, dy$$

By independence

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

Sum of random variables

Therefore

$$
\begin{aligned}
F_{X+Y}(a) &= \iint\limits_{x+y\leq a} f_X(x)\, f_Y(y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} f_Y(y) \left(\int_{-\infty}^{a-y} f_X(x)\, dx \right) dy \\
&= \int_{-\infty}^{\infty} f_Y(y) F_X(a-y)\, dy
\end{aligned}
$$

Sum of random variables

To compute PDF (probability density function), we differentiate equation (2) to get

$$
\begin{aligned}
f_{X+Y}(a) &= \frac{d}{da} \int_{-\infty}^{\infty} f_Y(y) F_X(a-y)\, dy \\
&= \int_{-\infty}^{\infty} \frac{d\, F_X(a-y)}{da}\, f_Y(y)\, dy \\
&= \int_{-\infty}^{\infty} f_X(a-y)\, f_Y(y)\, dy \qquad (2)
\end{aligned}
$$

Sum of random variables

Equation

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)\, f_Y(y)\, dy$$

is extremely useful and is often written using convolution symbol as

$$f_{X+Y} = f_X * f_Y$$