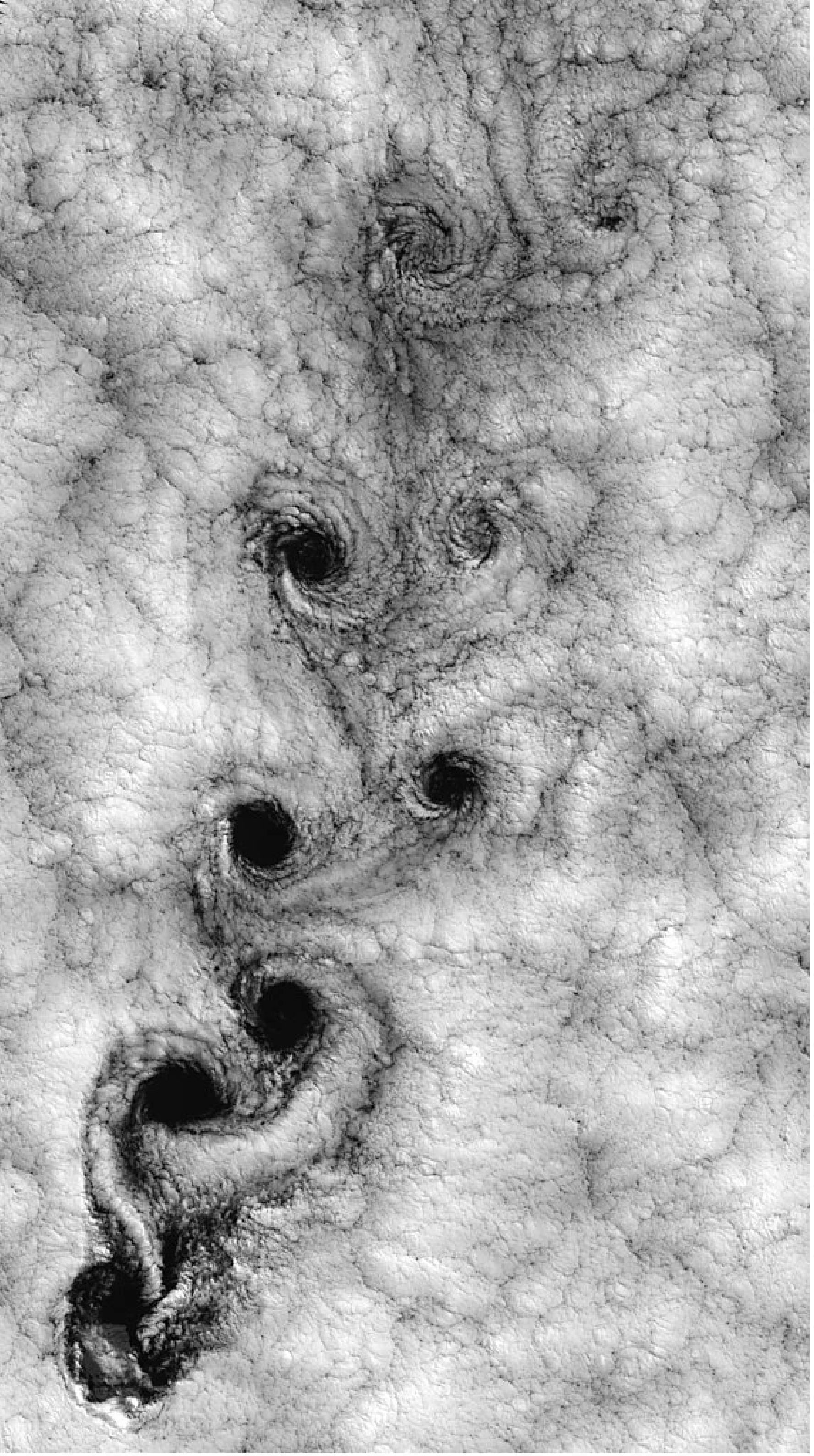

STATISTICS

LECTURE 4

Moment Generating Function

Moment Generating Function

Suppose X is a random variable. For a positive integer k,

$m_k = E[X^k]$ is known as the

k^{th} moment of X.

Moment Generating Function

Suppose X is a random variable. For a positive integer k,

$m_k = E[X^k]$ is known as the

k^{th} moment of X.

$$\sum_{x \in T} x^k P(X=x) \qquad \int_{-\infty}^{\infty} x^k f_X(x)\,dx.$$

Theorem Let X be a random variable. Let $k \in \mathbb{N}$. If $E[X^k] < \infty$ then $E[X^j] < \infty$ for $j \in \mathbb{N}$, $j \leq k$.

Prove by considering $|x| < 1$ and $|x| > 1$ separately.

Defn

Suppose X is a random variable.

and $D = \{t \in \mathbb{R} : E[e^{tX}] \text{ exists}\}$

The function $M : D \to \mathbb{R}$ given by

$M(t) = E[e^{tX}]$ is called the

moment generating function.

For a discrete r.v. $X: \Omega \to T$

$T = \{x_i : i \in \mathbb{N}\}$, for $t \in D$

$$\boxed{M_X(t) = \sum_{i \geq 1} e^{t x_i} P(X = x_i)}$$

whereas for a cts r.v.

$$\boxed{M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx}$$

Example 1

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

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} P(X=k)$$

$$= \sum_{k=0}^{\infty} \frac{e^{tk} e^{-\lambda} \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!}$$

$$= e^{-\lambda} e^{e^t \lambda} = e^{-\lambda(1-e^t)}$$

Example 2

$$X \sim \text{Geometric}(p)$$

$$M_X(t) = E[e^{tX}]$$

$$= \sum_{n=1}^{\infty} e^{tn} P(X=n)$$

$$= \sum_{n=1}^{\infty} e^{tn} p \cdot (1-p)^{n-1}$$

$$= pe^t \sum_{n=1}^{\infty} \left(e^t(1-p)\right)^{n-1}$$

$$= pe^t / 1 - e^t(1-p)$$

Example 3

$X \sim \text{Normal}(\mu, \sigma^2)$

$$M_X(t) = E[e^{tX}]$$

$$= \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{\sigma^2 \cdot 2}} \, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x^2 - (2\mu x + 2\sigma^2 t x) + \mu^2)/2\sigma^2} \, dx$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

Theorem Suppose for a r.v. X, $\exists\ \delta > 0$ s.t. $M_X(t)$ exists for $(-\delta, \delta)$.

(i) $$E[X^k] = M_X^{(k)}(0)$$

(ii) $$M_{aX}(t) = M_X(at), \text{ for } a \neq 0,\ at \in (-\delta, \delta)$$

(iii) $$M_{X+Y}(t) = M_X(t)\,M_Y(t)$$ for another independent r.v. Y whose m.g.f. also exists for $(-\delta, \delta)$.

"Proof"

$$M(t)$$

$$= E[e^{tX}]$$

$$= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} f(x)\,dx$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_{-\infty}^{\infty} x^k f(x)\,dx = \sum_{k=0}^{\infty} \frac{t^k}{k!} E[X^k]$$

"Proof"

$$M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} f(x)\,dx$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_{-\infty}^{\infty} x^k f(x)\,dx = \sum_{k=0}^{\infty} \frac{t^k}{k!} E[X^k]$$

$$\therefore \underline{M^{(k)}(0) = E[X^k]}$$

A valid proof would require advanced analysis which is not in the scope of this course.

Exercise

Use the previous theorem to find expected value and variance for Poisson, Geometric, Binomial, Normal distributions.

An important mathematical question was whether the knowledge of all the moments of a distribution uniquely determines that distribution.

An important mathematical question was whether the knowledge of all the moments of a distribution uniquely determines that distribution.

This was found to be false the two pdfs

$$f_1(x) = \frac{1}{\sqrt{2\pi}\,x} e^{-(\log x)^2/2} \qquad 0 \le x < \infty$$

and $f_2(x) = f_1(x)\left[1 + \sin(2\pi \log(x))\right], \; 0 \le x < \infty$

have the same moments.

Try!

A secondary question is whether the moment generating function exists provided all the moments are known.

This too may not be guaranteed. Outside of these counter-examples, we will focus on the situation where the moment generating function is convergent in a neighbourhood of zero.

Suppose X and Y are random variables.

The function

$M(s,t) = E[e^{sX+tY}]$ is called

the joint m.g.f. for X and Y.

Theorem

(I) Suppose X and Y are r.v.s and $M_X(t) = M_Y(t)$ in an open interval containing zero. Then X and Y have the same distribution.

(II) Suppose (X, W) and (Y, Z) are pairs of r.v.s and suppose $M_{X,W}(s,t) = M_{Y,Z}(s,t)$ in some rectangle around the origin, then (X,W) & (Y,Z) have the same joint distrib.

Theorem Suppose that (X, Y) are a pair of continuous r.v. with joint m.g.f. $M(s,t)$.

X and Y are independent if and only if

$$M(s,t) = M_X(s)\, M_Y(t).$$

<u>Proof</u> In one direction

$$M(s,t) = M(sX+tY)$$

$$= E[e^{sX+tY}]$$

$$= E[e^{sX} \cdot e^{tY}]$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{sx} e^{ty} f_X(x) f_Y(y)\, dx\, dy$$

$$= E[e^{sX}]\, E[e^{tY}] = M_X(s)\, M_Y(t)$$

Proof

In one direction

$$M(s,t) = M(sX+tY)$$
$$= E[e^{sX+tY}]$$
$$= E[e^{sX} \cdot e^{tY}]$$
$$= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{sx} e^{ty} f_X(x) f_Y(y)\, dx\, dy$$

Commuting variables in integration "Fubini"

$$= E[e^{sX}]\,E[e^{tY}] = M_X(s)\,M_Y(t)$$

Proof In the other direction,

If $M(s,t) = M_X(s) \cdot M_Y(t)$ then pick another pair $\hat{X}$ and $\hat{Y}$ that have the same distribution as X and Y but are independent.

Then by previous result

$$M_{\hat{X},\hat{Y}}(s,t) = M_{\hat{X}}(s)\, M_{\hat{Y}}(t)$$
$$= M_X(s)\, M_Y(t)$$
$$= M_{X,Y}(s,t).$$

By uniqueness of distributions under m.g.f., $(\hat{X},\hat{Y})$ and (X,Y) are identically distributed. $\square$

Example. let $X \sim \text{Normal}(\mu_1, \sigma_1^2)$ & $Y \sim \text{Normal}(\mu_2, \sigma_2^2)$ be independent.

then

$$M_X(at)M_Y(bt) = e^{a\mu_1 t + \frac{a^2\sigma_1^2 t^2}{2}} \times e^{b\mu_2 t + \frac{b^2\sigma_2^2 t^2}{2}}$$

$$= e^{(a\mu_1 + b\mu_2)t + \frac{a^2\sigma_1^2 + b^2\sigma_2^2}{2} t^2}$$

$\therefore \; aX + bY \sim \text{Normal}(a\mu_1 + b\mu_2, \; a^2\sigma_1^2 + b^2\sigma_2^2)$

Example. Let $X \sim \text{Normal}(\mu_1, \sigma_1^2)$ & $Y \sim \text{Normal}(\mu_2, \sigma_2^2)$ be independent.

then

$$M_X(at)\,M_Y(bt) = e^{a\mu_1 t + \frac{a^2\sigma_1^2 t^2}{2}} \times e^{b\mu_2 t + \frac{b^2\sigma_2^2 t^2}{2}}$$

$$= e^{(a\mu_1 + b\mu_2)t + \frac{a^2\sigma_1^2 + b^2\sigma_2^2}{2} t^2}$$

$\therefore$ $aX + bY \sim \text{Normal}(a\mu_1 + b\mu_2,\ a^2\sigma_1^2 + b^2\sigma_2^2)$

One can do this for n independent copies.

Exercises.

Let $X_1, X_2, \dots, X_n$ be i.i.d. rvs

Let $Y = X_1 + X_2 + \dots + X_n$. Prove that

$$\boxed{M_Y(t) = [M_{X_1}(t)]^n}$$

Let $Z = \dfrac{X_1 + X_2 + \dots + X_n}{n}$. Prove that

$$M_Z(t) = [M_{X_1}(t/n)]^n.$$

Samples and the double structure of a random variable

Samples and the double structure of a random variable

In studying probability theory, we have exact information about the random variable – we know its probability density.

With this information, we can derive a lot of characteristics – such as mean, variance etc.

In real life, we never know the exact distribution function. We can obtain estimates by

- taking a sample of observations,
- deriving an estimate.

In real life, we never know the exact distribution function. We can obtain estimates by

- taking a sample of observations,
- deriving an estimate.

We will assume that we have a random variable X and we take a sample on n observations.

The random variable has a dual nature.

Before the sample has been taken, the potential observations $\{X_1, X_2, \ldots, X_n\}$ may each be thought of as a copy of the random variable X. Each of them is itself a distinct random variable.

The random variable has a dual nature.

Before the sample has been taken, the potential observations $\{X_1, X_2, \ldots, X_n\}$ may each be thought of as a copy of the random variable X. Each of them is itself a distinct random variable.

Once the sample is obtained, we have n realizations of that random variable.

We will assume that $X_1, X_2, \dots, X_n$ are i.i.d. (independent and identically distributed)

The "empirical distribution" based on these is the discrete distribution with probability mass function given by

$$f_n(t) = \frac{1}{n} \left| \{ i : X_i = t \} \right| \quad \text{for } t \in \mathbb{R}$$

and the "empirical c.d.f."

$$F_n(x) = \frac{\left| \{ i : X_i \leq x \} \right|}{n}$$

Example Suppose that 10 people were surveyed and asked about how many litres of water they drink in a day

3 4 2 5 2 4 4 6 3 4

From this data, we can generate the empirical pmf

t	2	3	4	5	6	o/w
$f_{10}(t)$	$2/10$	$2/10$	$4/10$	$1/10$	$1/10$	0

This is a probability distribution made of realizations of the r.v. X.

Example

Suppose that 10 people were surveyed and asked about how many litres of water they drink in a day

3 4 2 5 2 4 4 6 3 4

From this data, we can generate the empirical pmf

t	2	3	4	5	6	o/w
$f_{10}(t)$	$2/10$	$2/10$	$4/10$	$1/10$	$1/10$	0

Sample mean Let $X_1, X_2, \ldots, X_n$ be i.i.d. r.v.'s with distribution X. The sample mean is also an r.v. given by $\bar{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$

For a sequence of fixed realizations of X, one obtains a realization of the r.v. $\bar{X}$.

On the other hand, the expected value μ of X is the "population mean".

Theorem Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample of random variables whose distribution has finite expected value μ and finite variance σ^2. Let $\bar{X}$ be the sample mean, then

$$E[\bar{X}] = \mu \quad \text{and} \quad SD[\bar{X}] = \frac{\sigma}{\sqrt{n}}.$$

Proof $$E[\bar{x}] = E\left[\frac{x_1 + x_2 + \cdots + x_n}{n}\right]$$

$$= \frac{1}{n}\left(E[x_1] + E[x_2] + \cdots + E[x_n]\right)$$

$$= \frac{n\mu}{n} = \mu.$$

& $$SD^2[\bar{x}] = Var\left[\frac{x_1 + x_2 + \cdots + x_n}{n}\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$\therefore$ $SD[\bar{x}] = \sigma/\sqrt{n}$. $\blacksquare$

Define the Sample Variance

$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n-1}$$

Define the Sample Variance

$$S^2 = \frac{(X_1-\bar{X})^2+(X_2-\bar{X})^2+\cdots+(X_n-\bar{X})^2}{n-1}$$

Theorem Let $X_1, X_2, \ldots, X_n$ be an i.i.d. Sample of random variables whose distribution has finite expectation μ and finite variance σ^2. Then

$$E[S^2] = \sigma^2$$

Proof

$$E[\bar{X}^2] = Var[\bar{X}] + E[\bar{X}]^2$$

$$= \frac{\sigma^2}{n} + \mu^2$$

and

$$E[X_j^2] = Var[X_j] + E[X_j]^2$$

$$= \sigma^2 + \mu^2$$

Proof

$$E[\bar{X}^2] = Var[\bar{X}] + E[\bar{X}]^2$$
$$= \frac{\sigma^2}{n} + \mu^2$$

and
$$E[X_j^2] = Var[X_j] + E[X_j]^2$$
$$= \sigma^2 + \mu^2$$

$$E[S^2] = \frac{1}{n-1}\left\{\sum_{j=1}^{n} E[(X_j - \bar{X})^2]\right\}$$
$$= \frac{1}{n-1}\left\{\sum_{j=1}^{n}\left(E(X_j^2) + E(\bar{X}^2) - 2E(X_j\bar{X})\right)\right\}$$
$$= \frac{1}{n-1}\left\{n E(X_j^2) - n E(\bar{X}^2)\right\}$$

Proof $E[\bar{X}^2] = Var[\bar{X}] + E[\bar{X}]^2$

$$= \frac{\sigma^2}{n} + \mu^2$$

and $E[X_j^2] = Var[X_j] + E[X_j]^2$

$$= \sigma^2 + \mu^2$$

use $\frac{\sum X_j}{n} = \bar{X}$

$$E[S^2] = \frac{1}{n-1}\left\{\sum_{j=1}^{n} E[(X_j - \bar{X})^2]\right\}$$

$$= \frac{1}{n-1}\left\{\sum_{j=1}^{n}\left(E(X_j^2) + E(\bar{X}^2) - 2E(X_j\bar{X})\right)\right\}$$

$$= \frac{1}{n-1}\left\{nE(X_j^2) - nE(\bar{X}^2)\right\}$$

Proof

$$E[\bar{x}^2] = Var[\bar{x}] + E[\bar{x}]^2$$
$$= \frac{\sigma^2}{n} + \mu^2$$

and
$$E[x_j^2] = Var[x_j] + E[x_j]^2$$
$$= \sigma^2 + \mu^2$$

use
$$\frac{\sum x_j}{n} = \bar{x}$$

$$E[S^2] = \frac{1}{n-1}\left\{\sum_{j=1}^{n} E[(x_j - \bar{x})^2]\right\}$$
$$= \frac{1}{n-1}\left\{\sum_{j=1}^{n}\left(E(x_j^2) + E(\bar{x}^2) - 2E(x_j\bar{x})\right)\right\}$$
$$= \frac{1}{n-1}\left\{n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2\right\} = \sigma^2$$

Sample Proportion

Theoretically, a distribution is known when $P(X \in A)$ is known for any event A.

However we can use the empirical distribution to estimate $P(X \in A)$

For this purpose, we define another variable

Theorem Let $X_1, X_2, \ldots, X_n$ be an i.i.d sample of r.v.s. with distribution X. Suppose that we are interested in $P(X \in A)$ for an event A.

Let
$$\hat{p}_n = \frac{|\{i : X_i \in A\}|}{n}$$

Then $E[\hat{p}_n] = P(X \in A)$ and

$$\operatorname{Var}(\hat{p}_n] \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Proof Let $Z_i = \begin{cases} 1 & \text{if } X_i \in A \\ 0 & \text{otherwise} \end{cases}$

then

$$\left|\{i : X_i \in A\}\right| = \sum_{i=1}^{n} Z_i$$

Z_i s are independent and identically distributed.

Proof Let $Z_i = \begin{cases} 1 & \text{if } X_i \in A \\ 0 & \text{otherwise} \end{cases}$

then

$$|\{i : X_i \in A\}| = \sum_{i=1}^{n} Z_i$$

Z_i s are independent and identically distributed.

$$P(Z_1 \in B_1, \dots, Z_n \in B_n)$$

$$= P(\mathbb{1}_A(X_1) \in B_1, \dots, \mathbb{1}_A(X_n) \in B_n)$$

$$= P(X_1 \in \mathbb{1}_A^{-1}(B_1), \dots, X_n \in \mathbb{1}_A^{-1}(B_n)) = \prod_{i=1}^{n} P(\mathbb{1}_A(X_i) \in B_i)$$

Proof Let $Z_i = \begin{cases} 1 & \text{if } X_i \in A \\ 0 & \text{otherwise} \end{cases}$

then

$$|\{i : X_i \in A\}| = \sum_{i=1}^{n} Z_i$$

Z_i s are independent and identically distributed.

where $Z_i \sim \text{Bernoulli}(p)$ where $p = P(X \in A)$

$\therefore$ $E[Z_i] = p$ $\quad Var(Z_i) = p(1-p)$

Therefore $\sum_{i=1}^{n} Z_i \sim \text{Binomial}(n, p)$

$\therefore$

$$E\left[\frac{\sum_{i=1}^{n} Z_i}{n}\right] = \frac{np}{n} = p$$

$$\text{Var}\left(\frac{\sum_{i=1}^{n} Z_i}{n}\right) = \frac{np(1-p)}{n^2}$$

$$= \frac{p(1-p)}{n} \longrightarrow 0$$

as $n \to \infty$

$\blacksquare$

Weak law of large numbers

Let $X_1, X_2, \ldots$ be a sequence of i.i.d r.vs. Assume that X_1 has finite mean μ and variance σ^2. Consider the sample mean of the first n X_i's

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n}, \quad \text{For any } \varepsilon > 0,$$

$$\lim_{n \to \infty} P\left(|\overline{X}_n - \mu| > \varepsilon\right) = 0.$$

Proof Each sample mean $\bar{X}_n$ has

$$E[\bar{X}_n] = \mu \quad \text{and} \quad Var[\bar{X}_n] = \frac{\sigma^2}{n}$$

$$\therefore \quad P(|\bar{X}_n - \mu| > \varepsilon)$$

$$= P\left((\bar{X}_n - \mu)^2 > \varepsilon^2\right) < \frac{\sigma^2}{\varepsilon^2 n} \quad \text{(by Chebyshev's inequality)}$$

This goes to zero as $n \to \infty$.

$\blacksquare$

This "notion" of convergence of

$$\bar{X}_n \xrightarrow{P} \mu$$

For any $\varepsilon > 0$,

$$\left(P\left(|\bar{X}_n - \mu| > \varepsilon\right) \to 0 \quad \text{as } n \to \infty \right)$$

is called <u>convergence in probability</u>

As a particular example, if an event "A" has probability P under some distribution X then if we count the no. of its occurences and divide by the no. of experiments this proportion $\hat{p}_n$ will converge to the probability $P(X \in A)$ by the weak law of large numbers. This shows the connection between frequency & probability.

Distributions derived from normal distributions

χ^2 distribution

If Z is a standard normal distribution, the distribution $U = Z^2$ is called the χ^2 distribution with 1 degree of freedom.

Let us compute the p.d.f. for χ^2.

Let us compute the p.d.f. for χ^2.

$$F_U(x) = P(U \leq x)$$

$$= P(-\sqrt{x} \leq Z \leq \sqrt{x})$$

$$= \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$$

where Φ is c.d.f. for Z.

Let us compute the p.d.f. for χ^2.

$$F_U(x) = P(U \le x)$$

$$= P(-\sqrt{x} \le Z \le \sqrt{x})$$

$$= \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$$

where Φ is c.d.f. for Z.

Differentiating w.r.t. x, we get

$$f_U(x) = \frac{1}{2} x^{-1/2} \phi(\sqrt{x}) + \frac{1}{2} x^{-1/2} \phi(-\sqrt{x})$$

$$= x^{-1/2} \phi(\sqrt{x}) = \frac{x^{-1/2} e^{-x/2}}{\sqrt{2\pi}}, \quad x \ge 0$$

If $U_1, U_2, \ldots, U_n$ are independent χ^2-distribution random variables with 1 degree of freedom, the random variable

$V = U_1 + U_2 + \cdots + U_n$ is called the χ^2 distribution with n degrees of freedom and written as χ^2_n

To compute the p.d.f. of χ_n^2, we need to know a little bit about the Gamma distribution and its p.d.f and m.g.f.

To compute the p.d.f of χ_n^2, we need to know a little bit about the Gamma distribution and its p.d.f and m.g.f.

The Gamma distribution depends on two parameters α and λ.

Its density is $$\boxed{g(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \quad t \geq 0}$$

where $\Gamma(x) = \int_0^\infty u^{x-1} e^{-u}\, du, \quad x > 0$

Let us find the m.g.f of gamma dist.

$$M(t) = \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\, dx$$

$$= \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x(\lambda-t)}\, dx$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha} = \left(\frac{\lambda}{\lambda-t}\right)^\alpha$$

$$, \ t<\lambda$$

Since m.g.f s of sums are products of m.g.f s for independent r.v.s.

(α_1, λ) (α_2, λ)

therefore Sum of gammas are also gammas $\sim (\alpha_1+\alpha_2, \lambda)$

Since m.g.f.s of sums are products of m.g.f.s for independent r.v.s.

therefore sum of $\underbrace{\text{gammas}}_{(\alpha_1, \lambda)\ (\alpha_2, \lambda)}$ are also $\underbrace{\text{gammas}}$ $\sim (\alpha_1 + \alpha_2, \lambda)$

Also notice that χ^2 is gamma distribution with parameters $1/2$ & $1/2$.

Therefore, χ_n^2 has gamma distribution with $\alpha = n/2$, $\lambda = 1/2$

and therefore the density is

$$f(v) = \frac{1}{2^{n/2}\,\Gamma(n/2)}\, v^{n/2-1} e^{-v/2}, \quad v \geq 0.$$

Therefore, χ_n^2 has gamma distribution with $\alpha = n/2$, $\lambda = 1/2$

and therefore the density is

$$f(v) = \frac{1}{2^{n/2}\,\Gamma(n/2)}\, v^{n/2 - 1} e^{-v/2}, \quad v \geq 0.$$

Its moment generating function is

$$M(t) = (1-2t)^{-n/2}$$

Defn A seq. $X_1, X_2, \ldots$ is said to converge in distribution to a random variable X if $F_{X_n}(x)$ converges to F_X at every point for which F_X is cts.

We write

$$X_n \xrightarrow{d} X$$

Example

$$X_n \sim \text{Uniform}(0, 1/n)$$

$$F_n(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ nx & \text{if } 0 < x < 1/n \\ 1 & \text{if } x \geq 1/n \end{cases}$$

The seq. of functions F_n converges to the function

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$