Chi-Square Dist is Normal Dist. Squared

Claim $\chi^2(1) \sim Z^2$, $Z \sim N(0,1)$

Chi-square w/ 1 df = Standard Normal Dist ($\mu=0, \sigma=1$)

Proof by CDF let $A \sim$ Distribution 0 (D_0) "Cumulative density function"
$B \sim D_1$

We wish to show $P(A<a) = P(B<b)$ $\forall a,b \in \mathbb{R}$

$\Rightarrow D_0 = D_1$

(1) $\frac{d}{dA}\left(P(A<a)\right) =$ "$f_A(a)$" derivative of cdf = pdf noted $f_X(x)$

(2) $\frac{d}{dB}\left(P(B<b)\right) = f_B(b)$

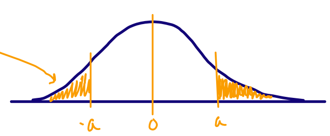
defining Pre-Req terms

(3) $Z \sim N(0,1) \Rightarrow$ "$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$" pdf of z

(4) $\Rightarrow F_Z(z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx$ cdf of z

(5) $V \sim \chi^2(1) \Rightarrow f_V(v) = \frac{e^{-v/2}}{\sqrt{2\pi v}}$ pdf of χ^2

(6) $F_Z(-a) = \Phi(-a)$ "cdf"



$\Rightarrow \Phi(-a) = 1 - \Phi(a)$

Assumption $\quad$ let $Z \sim N(0,1)$

Goal $\quad$ Prove: $Z^2 \sim \chi^2(1)$

We must show: $\quad$ By method of Distribution function

$$\iff P(Z^2 < z) = P(V < v) \quad \forall z, v \in \mathbb{R}, \quad V \sim \chi^2(1)$$

$$\iff \frac{d}{dz}\left(P(Z^2 < z)\right) = \frac{d}{dz}\left(P(V < v)\right)$$

RHS

$$\frac{d}{dv}\left(P(V < v)\right) = f_V(v)$$

$$= \frac{e^{-v/2}}{\sqrt{2\pi v}} \quad (5)$$

Note: $P(Z^2 < z) = P(-\sqrt{z} < \overset{\sim N(0,1)}{Z} < \sqrt{z})$

$$= \int_{-\sqrt{z}}^{\sqrt{z}} f_Z(x)\,dx$$

$$= \int_{-\infty}^{\sqrt{z}} f_Z(x)\,dx - \int_{-\infty}^{-\sqrt{z}} f_Z(x)\,dx$$

$$= \Phi(\sqrt{z}) - \Phi(-\sqrt{z}) \quad (4)$$

$$= \Phi(\sqrt{z}) - \left(1 - \Phi(\sqrt{z})\right) \quad (6)$$

$$P(Z^2 < z) = 2\,\Phi(\sqrt{z}) - 1 \quad (7)$$

LHS

$$\frac{d}{dz}\left(P(Z^2 < z)\right) = \frac{d}{dz}\left(2\,\Phi(\sqrt{z}) - 1\right) \quad (7)$$

$$= 2\,\frac{d}{dz}\left(\Phi(\sqrt{z})\right)$$

$$= 2\,\Phi'(\sqrt{z}) \cdot \frac{d}{dz}(\sqrt{z})$$

$$= 2\,\Phi'(\sqrt{z}) \cdot \frac{1}{2\sqrt{z}}$$

$$= 2\left[\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(\sqrt{z})^2}{2}}\right] \cdot \frac{1}{2\sqrt{z}}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{z}{2}} \cdot \frac{1}{\sqrt{z}}$$

$$= \frac{1}{\sqrt{2\pi z}} e^{-\frac{z}{2}}$$

let $z = v$ $\Rightarrow$ $\frac{1}{\sqrt{2\pi v}} e^{-\frac{z}{2}}$ which is our χ^2 pdf

Thus, RHS $\Rightarrow$ $Z^2 \sim \chi^2(1)$, $Z \sim N(0,1)$ ∎

$F = t^2$ when two-sample

Claim $F_{a-1, N-a} = \dfrac{MST}{MSE} = \dfrac{\frac{SST}{a-1}}{\frac{SSE}{N-a}}$ (1)

Reduces to

$$t_k^2 = \frac{(\bar{y}_1 - \bar{y}_2)^2}{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$ when $a = 2$ (2)

Symbol	Description
SSE	Sum of Squares due to Error
SST	Sum of Squares of Treatment
MSE	Mean Sum of squares Error
MST	Mean Sum of squares Treatment
a	Number of treatments
n_1	Number of observations in treatment 1
n_2	Number of observations in treatment 2
N	Total number of observations
$\bar{y}_{i.}$	Mean of treatment i
$\bar{y}_{..}$	Global mean
$k = N - a$	Degrees of freedom of the denominator of F

SSE is Sum of Squares (SS) $= SS_{y_1} + SS_{y_2}$

$SS = \text{variance}(n-1)$

Proof

Denominator of 1 $a=2$

$$MSE = \frac{SSE}{N-2} = \frac{\overbrace{\sum^{n_1} (y_{1i} - \bar{y}_1)^2}^{SS_{y_1}} + \overbrace{\sum^{n_2} (y_{2j} - \bar{y}_2)^2}^{SS_{y_2}}}{N-2}$$

$$S_i^2 = \frac{\sum^{n_i} (y_{ij} - \bar{y}_i)^2}{n_i - 1}$$ "sample variance estimator"

$$\Rightarrow MSE = \frac{SSE}{N-2} = S_p^2 = \boxed{\frac{S_1^2(n_1-1) + S_2^2(n_2-1)}{N-2}}$$

we'll call the pooled estimator

this variance

Numerator of 1 $a=2$

$$\frac{SST}{2-1} = SST$$

$$= \sum^{2} n_i (\bar{y}_i - \bar{y}_{..})^2$$

$$= \sum^{2} n_1 (\bar{y}_1 - \bar{y}_{..})^2 \quad \sum^{2} n_2 (\bar{y}_2 - \bar{y}_{..})^2$$

$$\bar{y}_{..} = \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}$$

$$\sum^{2} n_1 (\bar{y}_1 - \bar{y}_{..})^2 \quad \sum^{2} n_2 (\bar{y}_2 - \bar{y}_{..})^2$$

$$\boxed{SST = \underbrace{\sum^{2} n_1 \left(\bar{y}_1 - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2}_{\text{Part a}} + \underbrace{\sum^{2} n_2 \left(\bar{y}_2 - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2}_{\text{Part b}}}$$

Part a

$$n_1\left(\bar{y}_1 - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2$$

$$\Rightarrow n_1\left(\frac{N\bar{y}_1}{N} - \frac{n_1\bar{y}_1 + n_2\bar{y}_2}{N}\right)^2 = n_1\left(\frac{N\bar{y}_1 - n_1\bar{y}_1 - n_2\bar{y}_2}{N}\right)^2$$

$$= n_1\left(\frac{\bar{y}_1(N-n_1) - n_2\bar{y}_2}{N}\right)^2$$

$$= n_1\left(\frac{\bar{y}_1 n_2 - n_2\bar{y}_2}{N}\right)^2$$

$$= n_1\left(\frac{n_2(\bar{y}_1 - \bar{y}_2)}{N}\right)^2$$

$$= \boxed{\frac{n_1 n_2^2}{N^2}(\bar{y}_1 - \bar{y}_2)^2}$$

Part b

$$\vdots$$

$$= \boxed{\frac{n_2 n_1^2}{N^2}(\bar{y}_2 - \bar{y}_1)^2}$$

Numerator of 1 $a=2$

$$SST = \underbrace{\sum^{2}\frac{n_1 n_2^2}{N^2}(\bar{y}_1 - \bar{y}_2)^2}_{\text{Part a}} + \underbrace{\sum^{2}\frac{n_2 n_1^2}{N^2}(\bar{y}_2 - \bar{y}_1)^2}_{\text{Part b}}$$

$$\Rightarrow \sum^{2}\frac{n_1 n_2(n_1+n_2)}{N^2}(\bar{y}_2 - \bar{y}_1)^2$$

$$\Rightarrow \sum^{2}\frac{n_1 n_2 N}{N^2}(\bar{y}_2 - \bar{y}_1)^2$$

$$\Rightarrow \sum^{2}\frac{n_1 n_2}{N}(\bar{y}_2 - \bar{y}_1)^2$$

$$\boxed{SST = \sum^{2} \frac{1}{\frac{1}{n_1} + \frac{1}{n_2}} \left(\bar{y}_2 - \bar{y}_1 \right)^2}$$

All Together $a=2$

$$\boxed{SST = \sum^{2} \frac{1}{\frac{1}{n_1} + \frac{1}{n_2}} \left(\bar{y}_2 - \bar{y}_1 \right)^2}$$

$$\boxed{S_p^2 = \frac{S_1^2 (n_1 - 1) + S_2^2 (n_2 - 1)}{N - 2}}$$

$$F_{a-1, N-a} = \frac{MST}{MSE} = \frac{\frac{SST}{a-1}}{\frac{SSE}{N-a}} = \frac{\frac{SST}{2-1}}{\frac{SSE}{N-2}} = \frac{SST}{\frac{SSE}{N-2}} = \frac{SST}{S_p^2}$$

$$\Rightarrow \frac{\sum^{2} \frac{(\bar{y}_2 - \bar{y}_1)^2}{\frac{1}{n_1} + \frac{1}{n_2}}}{S_p^2} = \sum^{2} \frac{(\bar{y}_2 - \bar{y}_1)^2}{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = t_h^2 \quad \blacksquare$$

Central Limit Theorem

Claim

If random variables $X_1, \dots, X_n$ are Ind and identically distributed, with a constant fixed mean μ and constant finite variance σ^2, then the random variable Z approaches the standard Normal Dist. $N(0,1)$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

$$\mu = E(X_i) \qquad \sigma^2 = Var(X_i) \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad i = 1, \dots, n$$

Proof We define a Random Variable y_i that's Ind and Identically Dist.

$$y_i = \frac{X_i - \mu}{\sigma}$$

Thus,

$$E(y_i) = E\left(\frac{X_i - \mu}{\sigma}\right) = \frac{1}{\sigma} E(X_i - \mu)$$

$$= \frac{1}{\sigma}\left(E(X_i) - \mu\right)$$

$$= \frac{1}{\sigma}\left(\mu - \mu\right) = 0$$

$$Var(y_i) = Var\left(\frac{X_i - \mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(X_i - \mu)$$

$$= \frac{1}{\sigma^2} Var(X_i) = \frac{\sigma^2}{\sigma^2} = 1$$

We define a Random Variable $S = y_1 + \dots + y_n$ to be the sum of all y_i's

$$E(S) = E\left(\sum^{n} y_i\right) = \sum^{n} E(y_i) = 0$$

$$Var(S) = Var\left(\sum^{n} y_i\right) = \sum^{n} Var(y_i) = \sum^{n} 1 = 1 \cdot n = n$$

We define a Random Variable Z

$$Z = \frac{S\sqrt{n}}{n} = \frac{\sqrt{n}}{n}\sum^{n} y_i$$

$$= \frac{\sqrt{n}}{n}\sum^{n}\frac{X_i - \mu}{\sigma}$$

$$= \frac{\sqrt{n}}{n\sigma}\sum^{n}(X_i - \mu)$$

$$= \frac{\sqrt{n}}{n\sigma}\left[\sum^{n}(X_i) - n\mu\right]$$

$$= \frac{\sqrt{n}}{n\sigma}\sum^{n}(X_i) - \frac{\sqrt{n}}{n\sigma}\cdot n\mu$$

$$= \frac{\sqrt{n}}{\sigma}\bar{X} - \frac{\sqrt{n}\mu}{\sigma}$$

$$= \frac{\sqrt{n}}{\sigma}(\bar{X} - \mu)$$

$$Z = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$$

Moment generating functions

We first determine MGF of y_i

$$M_{y_i}(t) = 1 + \frac{t}{1!}E(y_i) + \frac{t^2}{2!}E(y_i^2) + \dots + \frac{t^n}{n!}E(y_i^n)$$

$$= 1 + \frac{t^2}{2!} + \dots + \frac{t^n}{n!}E(y_i^n)$$

since $E(y_i) = 0$
$E(y_i^2) = 1$

Then for MGF for $S = \sum^{n} y_i$

$$M_S(t) = \prod^{n} M_{y_i}(t) = \left(M_{y_i}(t)\right)^n$$

$$= \left(1 + \frac{t^2}{2!} + \dots + \frac{t^n}{n!}E(y_i^n)\right)^n$$

Then for MGF for $Z = \frac{S}{\sqrt{n}}$

$$m_z(t) = m_z\left(\frac{S}{\sqrt{n}}\right)$$

$$= \left(1 + \frac{t^2}{2!\,\sqrt{n}^2} + \dots + \frac{t^n}{n!\,\sqrt{n}^n} E(y_i^n)^n\right)^n$$

All together

$$\ln\left(m_z(t)\right) = \ln\left[\left(1 + \frac{t^2}{2!\,\sqrt{n}^2} + \dots + \frac{t^n}{n!\,\sqrt{n}^n} E(y_i^n)^n\right)^n\right]$$

$$= n \ln\left[\left(1 + \frac{t^2}{2!\,\sqrt{n}^2} + \dots + \frac{t^n}{n!\,\sqrt{n}^n} E(y_i^n)^n\right)\right]$$

Taylor/maclaurian series $\ln(1+x)$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \dots = \sum_{n=1}^{\infty} \frac{x^n}{n} (-1)^{n+1}$$

$$\ln\left(m_z(t)\right) = n \ln\left[1 + \frac{t^2}{2!\,\sqrt{n}^2} + \dots + \frac{t^n}{n!\,\sqrt{n}^n} E(y_i^n)^n\right]$$

$$= n \sum_{n=1}^{\infty} \frac{\left(\frac{t^2}{2!\,\sqrt{n}^2}\right)^n}{n} (-1)^{n+1}$$

$$\lim_{n\to\infty}\left[\ln\left(m_z(t)\right)\right] = \lim_{n\to\infty}\left[n \sum_{n=1}^{\infty} \frac{\left(\frac{t^2}{2!\,\sqrt{n}^2}\right)^n}{n} (-1)^{n+1}\right] = \frac{t^2}{2}$$

Finally

$$\ln\left(m_z(t)\right) = \frac{t^2}{2} \quad \text{as } n \to \infty$$

$$m_z(t) = e^{t^2/2}$$

Mathematical Proof: a logical set of steps that validates the truth of a general statement beyond any doubt.

- Hypothesis:
  - is always true
  - comes first in the statement (Denoted P)

- Notation:
  - " $\Rightarrow$ " implies
  - " $\in$ " is an element of
  - " $\equiv$ " is defined to be equal to
  - " $\exists$ " there exist
  - " $\nexists$ " there does not exist
  - " $\forall$ " for all

- Truth Table

P	Q	P => Q
T	T	T
T	F	F
F	T	T
F	F	T

Direct Proof: use theorems + Axioms to prove the conclusion of your Statement is true

Claim The sum of an odd and even positive integer is always odd.

Pf let a be even and b be odd.

$a = 2n$ and $b = 2m-1$ where $m, n \in \mathbb{Z}^+$

$a + b = 2n + 2m - 1$

$= 2(n+m) - 1$ which is odd since $n, m \in \mathbb{Z}^+$ ■

Claim The sum of two even numbers is always even

Pf let a, b be even.

$a = 2n$ and $b = 2m$, where $m, n \in \mathbb{Z}^+$

$a + b = 2n + 2m$

$= 2(n+m)$ which is even since $n, m \in \mathbb{Z}^+$ ■

2(stuff) which is even since $n, m \in \mathbb{Z}$. ■

Claim Show $\left(x + \left(\frac{a}{2}\right)\right)^2 - \left(\frac{a}{2}\right)^2 = x^2 + ax$

Pf
$$\left(x + \left(\frac{a}{2}\right)\right)^2 - \left(\frac{a}{2}\right)^2 = x^2 + ax + \frac{a^2}{4} - \frac{a^2}{4}$$
$$= x^2 + ax \quad ■$$

Proof by Contradiction : Assume $\neg Q$, Show contradiction of P

Contrapositive: Assume the second part of a statement is false and show that it leads to a contradiction of the hypothesis

If $P \Rightarrow Q$, then $\neg Q \Rightarrow \neg P$

Claim If integer n is odd, then n^2 is odd.

Pf Assume, to the contrary, n^2 is even
Then, $n^2 = 2b$ where $b \in \mathbb{Z}$

$\Rightarrow n \cdot n = 2b$ ⚡ Contradiction since we know the product of two even numbers must be even.

Thus, if n is odd n^2 must be odd. ■

Claim Show $\sqrt{2}$ is irrational

Pf Assume, to the contrary, $\sqrt{2}$ is rational.
Then, $\exists p, q \in \mathbb{Z}$ with $q \neq 0$ s.t. $\sqrt{2} = \frac{p}{q}$ irreducible

$\Rightarrow (\sqrt{2})^2 = \left(\frac{p}{q}\right)^2$

$\Rightarrow 2 = \frac{p^2}{q^2}$

$\Rightarrow p^2 = 2q^2$ Thus, p is even.

Then, $p = 2b$, $b \in \mathbb{Z}$

$\Rightarrow p^2 = 4b^2 = 2q^2$

$\Rightarrow 2b^2 = q^2$ Thus, q is even. ⚡ Contradiction

The quotient of two even numbers can be reduced.

Thus, $\sqrt{2}$ is irrational. ■

Claim] $\nexists\, x \in \mathbb{R}$ s.t. $\frac{1}{x-2} = 1 - x$

Pf] Assume, to the contrary, $\exists\, a \in \mathbb{R}$ s.t. $\frac{1}{a-2} = 1 - a$

Then,

$$1 = (a-2)(1-a)$$
$$\Rightarrow 1 = a - a^2 - 2 + 2a$$
$$\Rightarrow 0 = -a^2 + 3a - 3$$
$$\Rightarrow a^2 - 3a + 3 = 0$$
$$\Rightarrow a = \frac{3 \pm \sqrt{9 - 12}}{2} \notin \mathbb{R} \quad \text{※ Contradiction}$$

Thus, $\nexists\, a \in \mathbb{R}$ s.t. $\frac{1}{a-2} = 1 - a$

Counterexample: an acceptable proof that a statement is false

Claim] If $n \in \mathbb{Z}$ and n^2 is divisible by 4, then n is divisible by 4

Pf] let $n = 6$

Then,

$$n^2 = 36 / 6 \checkmark$$

But $n = 6$ is not divisible by 6 ■

n	n^2
2	4
$\cancel{3}$	$\cancel{4}$
$\cancel{4}$	$\cancel{16}$
$\cancel{5}$	$\cancel{25}$
6	36

Proof by Induction

Skeleton:

1. $P(n)$
2. Show $P(1)$ true "basic step"
3. Assume $P(k)$ true $\forall k \in \{\text{set}\}$
4. Show $P(k+1)$ true using the Hypothesis (Assumption from ③)
5. Thus, by Principal of Mathematical Induction ....

Claim] $1 + 2 + 3 + \cdots + (n-1) + n + (n-1) + \cdots + 3 + 2 + 1 = n^2$

Pf] $P(n)$: $1 + 2 + 3 + \cdots + (n-1) + n + (n-1) + \cdots + 3 + 2 + 1 = n^2$

$P(1)$: $1 = 1^2 \checkmark$

Assume $P(k)$ true $\forall k \in \mathbb{Z}^+$:

$$1+2+3+\cdots+(k-1)+k+(k-1)+\cdots+3+2+1=k^2$$

Want to show $P(k+1)$ true:

$$1+2+3+\cdots+(k-1)+k+(k+1)+k+(k-1)+\cdots+3+2+1$$

$$=1+2+3+\cdots+(k-1)+k+(k-1)+\cdots+3+2+1+(k+1)+k \quad \text{Rearrange}$$

$$=k^2+(k+1)+k \quad \text{by Hypothesis}$$

$$=k^2+2k+1$$

$$=(k+1)^2 \checkmark$$

Thus, since k is arbitrary $P(k+1)$ true $\forall k \in \mathbb{Z}^+$
by PMI $P(n)$ true ∎

Claim: $3^{2n}+7$ is divisible by 8 $\forall n \in \mathbb{N}$

Pf: $P(n) = 3^{2n}+7 \backslash 8$

$P(0) = 3^0+7 = 8 \checkmark$

Assume $P(k)$ true $\forall k \in \mathbb{N}$: $3^{2k}+7 \backslash 8$

Thus, $\exists a \in \mathbb{Z}^+$ s.t. $3^{2k}+7=8a$

Want to show $P(k+1)$ true: $3^{2(k+1)}+7 \backslash 8$

$$=3^{2k+2}+7$$

$$=3^{2k}\cdot 9+7$$

$$=(8a-7)\cdot 9+7$$

$$=72a-63+7$$

$$=72a-56$$

$$=8(9a-7) \checkmark$$

Thus, since k is arbitrary $P(k+1)$ true $\forall k \in \mathbb{Z}^+$
by PMI $3^{2n}+7$ is divisible by 8 $\forall n \in \mathbb{N}$ ∎