III-2-5. Chai square distribution

In normal distribution, the probability of variable (x) is as follow

$$\mathrm{P}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

In this formula, the meaning of $\frac{x-\mu}{\sigma}$ is standardization of distance from mean by standard deviation of parent population. A result of the standardization, we can treat the distribution as standard normal distribution and we can discuss the possibility of obtained data. Chai square is square of standardized distance. $\left(\frac{x-\mu}{\sigma}\right)^2$

$$\chi^2 = \left(\frac{x-\mu}{\sigma}\right)^2$$

When we consider standardized distance is x,

$$\chi^2 = x^2$$

Some may consider that is better to name $\left(\frac{x-\mu}{\sigma}\right)^2$ as x^2 (square of ex).

However, when we use χ^2, χ^2 is not square of χ, but we are saying that we will treat χ^2 one single variable. We use χ^2 instead of x^2 in this meaning, and when $\mathrm{P}(x)$ shows normal distribution, $\mathrm{P}(\chi^2)$ shows χ^2 distribution.

Chai square analysis is commonly used for judgement of unlikeliness. When value of χ^2 is larger than a definite value which is determined depending on a risk level (p),we can conclude that the data is stochastically not obtainable presuming that the expected mean is correct. Inversely, we can say that the expectated mean is not correct in the risk level, when larger χ^2 value is obtained. More simply, χ^2 analysis is used to say that the expectation is not correct, and it cannot be used as warrant for correctness of expectation. Most interesting nature of χ^2 distribution is reproductive property. Meaning of reproductive property is nature that the sum of two variance obey the same distribution family of original variances. In this case, when two variables fluctuate in χ^2 distribution the variable made by summing of the two variables fluctuate in χ^2 distribution.

$$\mathrm{X}_1 \sim {\chi^2}_m, \mathrm{X}_2 \sim {\chi^2}_n \rightarrow \mathrm{X}_1 + \mathrm{X}_2 \sim {\chi^2}_{m+n}$$

Here, ${\chi^2}_n$ means χ^2 of which degree of freedom is n. This nature is important, because obtainable expectation value is generally more than two.

Using nature of chai square distribution, we can calculate chai square distribution as follow.

$${\chi^2}_{\phi:Z_1,\cdots,Z_\phi} = Z_1 + \cdots + Z_\phi$$
$${\chi^2}_{\phi-1:Z_1,\cdots,Z_{\phi-1},} = Z_1 + \cdots + Z_{\phi-1}$$

$$\chi^2{}_{1:Z_\phi} = Z_\phi$$
$$Z_\phi: stachastic\ event\ follwoing\ normal\ distribution$$

We consider $\chi^2{}_n$ is a variable. We make probability distribution of $\chi^2{}_n$ from following standard normal distribution

$$P(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2}$$

Cumulative probability from 0 to x is as follow.

$$\phi(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2} dx$$

Standard normal distribution is symmetric centering 0 and $x^2 = (-x)^2$. The distribution of probability $P(x^2)$ is double in height and expand horizontally in square dimension as piling up after doubling up at center (see Fig.18)

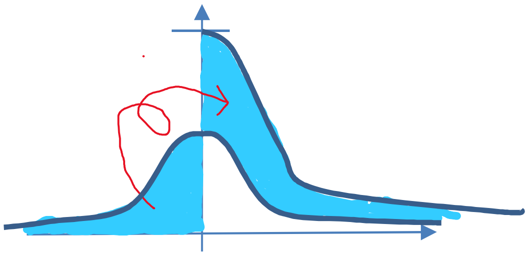


Fig. 18. Superposition of probability after half fold

$$\phi(x^2) = 2\int_0^x \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2} dx$$

Probability is derivative value of cumulative probability

$$P(x) = \frac{d\phi(x)}{dx}$$

$$P(\chi^2{}_1) = \frac{d\phi(x)}{d\chi^2{}_0}$$

$$\chi^2{}_1 = x^2$$

$$\frac{d\chi^2{}_1}{dx} = 2x = 2\sqrt{\chi^2{}_1}$$

$$P(\chi^2{}_1) = 2\frac{d\phi(x)}{d\chi^2{}_1} = 2\frac{d\phi(x)}{dx}\frac{dx}{d\chi^2{}_1} = 2P(x)\frac{1}{2\sqrt{\chi^2{}_1}} = \frac{1}{\sqrt{2\pi\chi^2{}_1}} e^{\frac{-\chi^2{}_1}{2}}$$

$$P(\chi^2{}_1) = \frac{1}{\sqrt{2\pi\chi^2{}_1}} e^{\frac{-\chi^2{}_1}{2}}$$

When we sum up all possible probability in standard normal distribution, the result should be 1.

$$\int_{-\infty}^{\infty} p(x)dx = 2\int_{0}^{\infty} p(x)dx = 1$$

When phenomenon x and phenomenon y is independent and $x \sim N(0,1)$ and $y \sim N(0,1)$.

The probability of happening phenomenon x and phenomenon y simultaneously is obtained by multiplication of both probabilities.

$$P(x, y) = P(x)P(y)$$

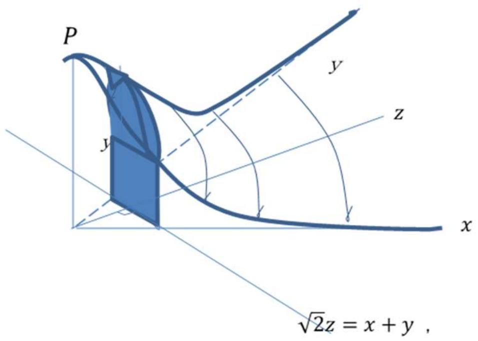
We consider two chai squares. One is ${\chi^2}_x$, the other is ${\chi^2}_y$ and consider the probability of sum of ${\chi^2}_{1:x}$ and ${\chi^2}_{1:y}$ as a certain value and express it as ${\chi^2}_{2:x+y}$

What we do in this paragraph is formulization of $\mathrm{P}\left({\chi^2}_{2:x+y}\right)$

When we express the sum of total probability by definite multiple integral from 0 to ∞.

$$\int_{0}^{\infty} P({\chi^2}_{2:x+y})d{\chi^2}_{2:x+y} = 2\int_{0}^{\infty}\int_{0}^{\infty} P({\chi^2}_{1:x})P({\chi^2}_{1:y})d{\chi^2}_{1:x}d\,{\chi^2}_{1:y} = 1$$

It is difficult to calculate integral in this coordinate, and we consider $-\frac{\pi}{4}$ rotation of coordinate. Because ${\chi^2}_{2:x+y}$ becomes parallel line of new axis $(w)$and we can integrate the cross section on the line to orthogonal orientation (z)



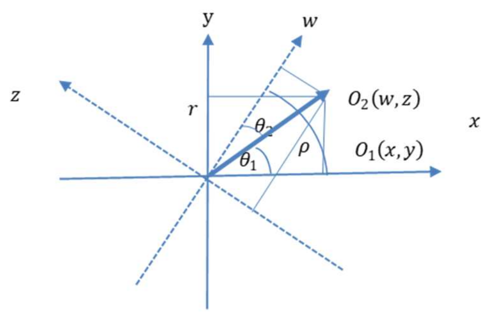


Fig. 19. Transformation of coordinate (rotation of axes)

We transform orthogonal coordinate, $O_1(x, y)$, to polar coordinate $P_1(r, \theta_1)$
Then rotate the axes of $P_1(r, \theta_1)$ for transformation to $P_2(r, \theta_2)$
Then transform $P_2(r, \theta_2)$ to $O_2(w, z)$. (See III-3-3. Jacobian and III-3-4. Polar coordinate)

$O_1(x, y) \to P_1(r, \theta_1)$

$$x = r\cos\theta_1$$

$$y = r\sin\theta_1$$

$O_2(w, z) \to P_2(r, \theta_2)$

$$w = r\cos\theta_2$$

$$z = r\sin\theta_2$$

$\theta_2 = \theta_1 - \rho$

$$w = r\cos(\rho - \theta_1) = r(\cos\rho\cos\theta_1 + \sin\rho\sin\theta_1) = x\cos\rho + y\sin\rho$$

$$z = r\sin(\rho - \theta_1) = r(\sin\rho\cos\theta_1 - \cos\rho\sin\theta_1) = y\cos\rho - x\sin\rho$$

Here, $\rho = -\frac{\pi}{4}$

$$\sin\rho = \sin\left(-\frac{\pi}{4}\right) = -\sin\frac{\pi}{4} = -\frac{1}{\sqrt{2}}$$

$$\cos\rho = \cos\left(-\frac{\pi}{4}\right) = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$$

$$w = \frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y$$

$$z = \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y$$

$$x = \frac{1}{\sqrt{2}}w + \frac{1}{\sqrt{2}}z$$

$$y = -\frac{1}{\sqrt{2}}w + \frac{1}{\sqrt{2}}z$$

$$\frac{dx}{dz} = \frac{1}{\sqrt{2}}, \qquad \frac{dx}{dw} = \frac{1}{\sqrt{2}}$$

$$\frac{dy}{dz} = \frac{1}{\sqrt{2}}, \qquad \frac{dy}{dw} = -\frac{1}{\sqrt{2}}$$

When we express the sum of total probability by infinite multiple integral.

$$\int_0^\infty P(\chi^2{}_{2:x+y})d\chi^2{}_{2:x+y} = 2\int_0^\infty\int_0^\infty P(\chi^2{}_{1:x})P(\chi^2{}_{1:y})d\chi^2{}_{1:x}d\chi^2{}_{1:y} = 1$$

About $\int\int P(\chi^2{}_{1:x})P(\chi^2{}_{1:y})d\chi^2{}_{1:x}d\chi^2{}_{1:y}$,

$$P(\chi^2{}_{1:x}) = \frac{1}{\sqrt{2\pi\chi^2{}_{1:x}}} e^{\frac{-\chi^2{}_{1:x}}{2}}$$

$$P(\chi^2{}_{1:y}) = \frac{1}{\sqrt{2\pi\chi^2{}_{1:y}}} e^{\frac{-\chi^2{}_{1:y}}{2}}$$

$$\int\int P(\chi^2{}_{1:x}) P\left(\chi^2{}_{1:y}\right) d\chi^2{}_{1:x}\, d\chi^2{}_{1:y}$$

$$= \int\int \frac{1}{\sqrt{2\pi\chi^2{}_{1:x}}} e^{\frac{-\chi^2{}_{1:x}}{2}} \cdot \frac{1}{\sqrt{2\pi\chi^2{}_{1:y}}} e^{\frac{-\chi^2{}_{1:y}}{2}} d\chi^2{}_{1:x} d\chi^2{}_{1:y}$$

$$= \frac{1}{2\pi}\int\int \frac{1}{\sqrt{\chi^2{}_{1:x}\chi^2{}_{1:y}}} e^{\frac{-\left(\chi^2{}_{1:x}+\chi^2{}_{1:y}\right)}{2}} d\chi^2{}_{1:x} d\chi^2{}_{1:y}$$

$$= \frac{1}{2\pi}\int\int \frac{1}{\sqrt{\chi^2{}_{1:x}\chi^2{}_{1:y}}} e^{\frac{-\chi^2{}_{2:x+y}}{2}} d\chi^2{}_{1:x} d\chi^2{}_{1:y}$$

Our purpose is to make formula of cumulative volume from 0 to z. This can be obtained by integral of cross section surface on line $y = \sqrt{2}z - x$ (parallel lines of w on $w - z$ plane) along z. Area of cross section surface is obtainable by integral of probability y on the parallel line along w from z to -z.

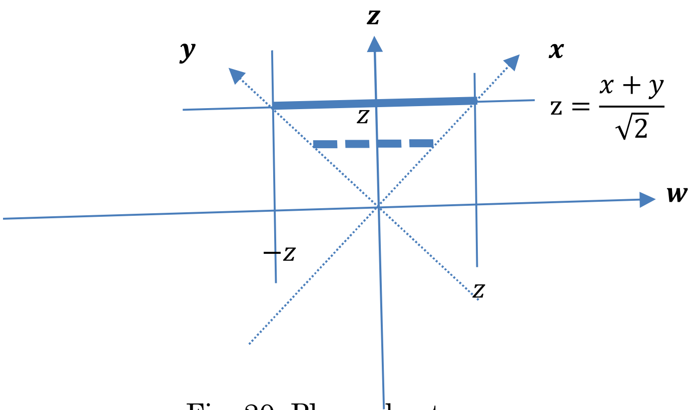


Fig. 20. Plane chart on $w - z$

Transformation

$$z = \frac{1}{\sqrt{2}}(x + y) = \frac{1}{\sqrt{2}}\chi^2{}_{2:x+y}$$

$$x = \frac{1}{\sqrt{2}}w + \frac{1}{\sqrt{2}}z$$

$$y = -\frac{1}{\sqrt{2}}w + \frac{1}{\sqrt{2}}z$$

$$xy = \frac{1}{2}(z^2 - w^2)$$

$$\frac{dx}{dz} = \frac{1}{\sqrt{2}}, \quad \frac{dy}{dw} = -\frac{1}{\sqrt{2}}$$

$$dx = \frac{1}{\sqrt{2}}dz, \quad dy = -\frac{1}{\sqrt{2}}dw$$

$$\int_0^\infty \int_0^\infty p(x)P(y)dxdy = \frac{1}{2\pi}\int_0^\infty \int_0^\infty \frac{1}{\sqrt{xy}} e^{\frac{-\left(\chi^2{}_{1:x}+\chi^2{}_{1:y}\right)}{2}} d\chi^2{}_{1:x} d\chi^2{}_{1:y}$$

$$= \frac{1}{2\pi}\int\int_D \frac{1}{\sqrt{\frac{1}{2}(z^2 - w^2)}} e^{\frac{-\sqrt{2}z}{2}} \frac{1}{2} d\chi^2{}_{1:x} d\chi^2{}_{1:y}$$

When we give domain for definite integral,

$$\mathrm{D} = \{(w, z) | -z \le w \le z, \ \ 0 \le z \le \infty\}$$

$$\int_0^\infty \int_0^\infty p(x)P(y)dxdy = \frac{1}{2\sqrt{2}\pi}\int_0^\infty \int_{-z}^{z} \frac{1}{\sqrt{(z^2 - w^2)}} e^{\frac{-z}{\sqrt{2}}} dwdz$$

$$= \frac{1}{2\sqrt{2}\pi}\int_0^\infty \int_{-z}^{z} \frac{1}{z\sqrt{\left(1 - \left(\frac{w}{z}\right)^2\right)}} e^{\frac{-z}{\sqrt{2}}} dwdz$$

$$= \frac{1}{2\sqrt{2}\pi}\int_0^{z_l} \frac{1}{z} e^{\frac{-z}{\sqrt{2}}} \left(\int_{-1}^{1} \frac{1}{\sqrt{(1 - u^2)}} du\right) dz$$

$$\because \ \frac{w}{z} = u, \qquad \frac{1}{z} = \frac{du}{dw}$$

$$\int_{-z}^{z} \frac{1}{\sqrt{(z^2 - w^2)}} dw = \int_{-1}^{1} \frac{1}{\sqrt{(1 - (u)^2)}} du$$

$$\int_{-1}^{1} \frac{1}{\sqrt{(1 - (u)^2)}} du = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta = [\theta]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \pi$$

$$\because \ u = \sin\theta\,, \qquad \frac{du}{d\theta} = \cos\theta$$

$$(1 - (u)^2) = 1 - \sin^2\theta = \cos^2\theta$$

$$\frac{1}{\sqrt{(1 - (u)^2)}} = \frac{1}{\cos\theta}$$

Conclusively,

$$\int_0^\infty \int_0^\infty p(x)P(y)dxdy = \frac{1}{2\sqrt{2}\pi}\int_0^\infty \int_{-z}^{z} \frac{1}{\sqrt{(z^2-w^2)}} e^{\frac{-z}{\sqrt{2}}} dwdz$$

$$= \frac{1}{2\sqrt{2}\pi}\int_0^\infty \pi e^{\frac{-z}{\sqrt{2}}} dz = \frac{1}{2\sqrt{2}}\int_0^\infty e^{\frac{-z}{\sqrt{2}}} dz$$

$$= \frac{1}{2\sqrt{2}}\int_0^\infty e^{\frac{-{\chi^2}_{2:x+y}}{2}} \frac{1}{\sqrt{2}} d{\chi^2}_{2:x+y} = \frac{1}{4}\int_0^\infty e^{\frac{-{\chi^2}_{2:x+y}}{2}} d{\chi^2}_{2:x+y}$$

$$= \frac{1}{4}\int_0^\infty e^{-t} 2dt = -\frac{1}{2}[e^{-t}]_0^\infty = \frac{1}{2}(1-e^{-\infty}) = \frac{1}{2}$$

$$\because z = \frac{1}{\sqrt{2}}{\chi^2}_{2:x+y}$$

$$\frac{dz}{d{\chi^2}_{2:x+y}} = \frac{1}{\sqrt{2}}$$

$$\mathrm{t} = \frac{{\chi^2}_{2:x+y}}{2}$$

$$\frac{dt}{d{\chi^2}_{2:x+y}} = \frac{1}{2}$$

We could confirm

$$\int_0^\infty P({\chi^2}_{2:x+y}) d{\chi^2}_{2:x+y} = 1$$

and cumulative probability from 0 to ${\chi^2}_{2:x+y}$ is

$$\phi\left({\chi^2}_{2:x+y}\right) = \frac{1}{2}\int_0^{{\chi^2}_{2:x+y}} e^{\frac{-{\chi^2}_{2:x+y}}{2}} d{\chi^2}_{2:x+y}$$

We can obtain $\mathrm{P}\left({\chi^2}_{2:x+y}\right)$ by differential of $\phi\left({\chi^2}_{2:x+y}\right)$

$$\mathrm{P}\left({\chi^2}_{2:x+y}\right) = \frac{1}{2} e^{\frac{-{\chi^2}_{2:x+y}}{2}}$$

We could formularize $\mathrm{P}\left({\chi^2}_{2}\right)$ from $\mathrm{P}\left({\chi^2}_{1}\right)$. This result means that we can formularize $P\left({\chi^2}_{\phi}\right)$ by repeating similar procedure. However, it is time consuming, and there more generalized expression of the formula include ϕ in the formula. Before that the author explains the meaning of ϕ. The name of ϕ is degree of freedom. Concept of degree of freedom often appear in statistical analysis. Basic mean of degree of freedom is number of elements which can vary in a constraint condition. As an example, assuming that there are red and white balls and total number of balls are 10. Number of red balls

fluctuates from 0 to10. When the number of red balls is fixed at a number, number of white balls are automatically determined. Inversely, when we fix the number of white balls, the number of red balls are automatically determined. In this case only one element fluctuate freely in the constraint condition and degree of freedom is 1. When there are three colors of ball, red, white and green, the degree of freedom is 2.
Using degree of freedom, our results obtained is as follow.
When $\phi = 1$,

$$\mathrm{P}({\chi^2}_1) = \frac{1}{\sqrt{2\pi {\chi^2}_1}} e^{\frac{-{\chi^2}_1}{2}}$$

When $\phi = 2$,

$$\mathrm{P}({\chi^2}_2) = \frac{1}{2} e^{-\frac{1}{2}{\chi^2}_2}$$

When $\phi = 3$
We repeat the same procedure of making $\mathrm{P}({\chi^2}_2)$ from $\mathrm{P}({\chi^2}_1)$.

$$\int_0^{\infty} P({\chi^2}_3) d{\chi^2}_3 = 2\int_0^{\infty}\int_0^{\infty} P({\chi^2}_1)P({\chi^2}_2) d{\chi^2}_2 d{\chi^2}_1 = 1$$

About $\int\int P({\chi^2}_1)P({\chi^2}_2)d{\chi^2}_2 d{\chi^2}_1$,

$$P({\chi^2}_1) = \frac{1}{\sqrt{2\pi {\chi^2}_1}} e^{\frac{-{\chi^2}_1}{2}}$$

$$\mathrm{P}({\chi^2}_2) = \frac{1}{2} e^{-\frac{1}{2}{\chi^2}_2}$$

$$\int\int P({\chi^2}_1)P({\chi^2}_2)d{\chi^2}_2 d{\chi^2}_1$$

$$= \int\int \frac{1}{\sqrt{2\pi {\chi^2}_1}} e^{\frac{-{\chi^2}_1}{2}} \cdot \frac{1}{2} e^{\frac{-{\chi^2}_2}{2}} d{\chi^2}_2 d{\chi^2}_1$$

$$= \frac{1}{2\sqrt{2\pi}} \int\int \frac{1}{\sqrt{{\chi^2}_1}} e^{\frac{-({\chi^2}_1 + {\chi^2}_2)}{2}} d{\chi^2}_2 d{\chi^2}_1$$

$$= \frac{1}{2\sqrt{2\pi}} \int\int \frac{1}{\sqrt{{\chi^2}_3 - {\chi^2}_2}} e^{\frac{-{\chi^2}_3}{2}} d{\chi^2}_2 d{\chi^2}_3$$

$$\because {\chi^2}_1 + {\chi^2}_2 = {\chi^2}_3$$

$$1=\frac{d\chi^2{}_3}{d\chi^2{}_1}$$

Here, $\chi^2{}_2=u\chi^2{}_3$

$$\chi^2{}_3-\chi^2{}_2=\chi^2{}_3(1-u)$$

$$0\leq u\leq 1$$

$$\frac{d\chi^2{}_2}{du}=\chi^2{}_3$$

$$\int\int P(\chi^2{}_1)P(\chi^2{}_2)d\chi^2{}_2\,d\chi^2{}_1$$

$$=\frac{1}{2\sqrt{2\pi}}\int\int\frac{1}{\sqrt{\chi^2{}_3(1-u)}}e^{\frac{-\chi^2{}_3}{2}}\chi^2{}_3dud\chi^2{}_3$$

$$=\frac{1}{2\sqrt{2\pi}}\int\int\frac{\chi^2{}_3}{\sqrt{\chi^2{}_3}}\frac{1}{\sqrt{(1-u)}}e^{\frac{-\chi^2{}_3}{2}}dud\chi^2{}_3$$

$$\frac{1}{2\sqrt{2\pi}}\int\frac{\chi^2{}_3}{\sqrt{\chi^2{}_3}}e^{\frac{-\chi^2{}_3}{2}}\left(\int_0^1\frac{1}{\sqrt{(1-u)}}du\right)d\chi^2{}_3$$

$$\frac{1}{2\sqrt{2\pi}}\int\chi^2{}_3{}^{\frac{1}{2}}e^{\frac{-\chi^2{}_3}{2}}\left[-2\sqrt{(1-u)}\right]_0^1d\chi^2{}_3$$

$$\frac{1}{2\sqrt{2\pi}}\int\chi^2{}_3{}^{\frac{1}{2}}e^{\frac{-\chi^2{}_3}{2}}\left(-2\sqrt{1-1}+2\sqrt{1-0}\right)d\chi^2{}_3$$

$$=\frac{1}{\sqrt{2\pi}}\int\chi^2{}_3{}^{\frac{1}{2}}e^{\frac{-\chi^2{}_3}{2}}d\chi^2{}_3$$

$$\Phi(\chi^2{}_3)=\frac{1}{\sqrt{2\pi}}\int_0^{\chi^2{}_3}\chi^2{}_3{}^{\frac{1}{2}}e^{\frac{-\chi^2{}_3}{2}}d\chi^2{}_3$$

$$\mathrm{P}(\chi^2{}_3)=\frac{d\Phi(\chi^2{}_3)}{d\chi^2{}_3}=\frac{d\Phi(\chi^2{}_3)}{dt}\frac{dt}{d\chi^2{}_3}=\frac{1}{\sqrt{2\pi}}(2t)^{\frac{1}{2}}e^{-t}2dt=\frac{\sqrt{2}}{\sqrt{\pi}}\chi^2{}_3{}^{\frac{1}{2}}e^{\frac{-\chi^2{}_3}{2}}$$

$$\because\frac{\chi^2{}_3}{2}=t$$

$$d\chi^2{}_3=2dt$$

$$\mathrm{P}(\chi^2{}_3)=\frac{\sqrt{2}}{\sqrt{\pi}}\chi^2{}_3{}^{\frac{1}{2}}e^{\frac{-\chi^2{}_3}{2}}$$

When we compare $\mathrm{P}(\chi^2{}_1),\mathrm{P}(\chi^2{}_2),\mathrm{P}(\chi^2{}_3)$

$$P(\chi^2{}_1) = \frac{1}{\sqrt{2\pi\chi^2{}_1}} e^{\frac{-\chi^2{}_1}{2}}$$

$$\mathrm{P}(\chi^2{}_2) = \frac{1}{2} e^{-\frac{1}{2}\chi^2{}_2}$$

$$\mathrm{P}(\chi^2{}_3) = \frac{\sqrt{2}}{\sqrt{\pi}} \chi^2{}_3{}^{\frac{1}{2}} e^{\frac{-\chi^2{}_3}{2}}$$

We can consider possibility of a rule. That is

$$\mathrm{P}(\chi^2{}_\phi) = A_\phi \left(\chi^2{}_\phi\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}}$$

We confirm adequacy of this hypothesis.

$$\mathrm{P}(\chi^2{}_\phi) = A_\phi \left(\chi^2{}_\phi\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}}$$

When the hypothesis is correct

$$\mathrm{P}(\chi^2{}_1) = A_1 (\chi^2{}_1)^{\frac{1}{2}-1} e^{\frac{-\chi^2{}_1}{2}}$$

$$\mathrm{P}(\chi^2{}_{\phi-1}) = A_{\phi-1} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-1}{2}-1} e^{\frac{-\chi^2{}_{\phi-1}}{2}}$$

$$\mathrm{P}(\chi^2{}_\phi) = A_\phi \left(\chi^2{}_\phi\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}}$$

When we consider $\chi^2\phi_{-1}$fluctuate from 0 to $\chi^2{}_\phi$

$$\mathrm{P}\left(\chi^2{}_\phi\right) = \int_0^{\chi^2\phi} \mathrm{P}(\chi^2{}_1)\mathrm{P}(\chi^2{}_{\phi-1})\, d\chi^2{}_{\phi-1}$$

$$A_\phi \left(\chi^2{}_\phi\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}} = \int_0^{\chi^2\phi} A_{\phi-1} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-1}{2}-1} e^{\frac{-\chi^2{}_{\phi-1}}{2}} \cdot A_1 (\chi^2{}_1)^{\frac{1}{2}-1} e^{\frac{-\chi^2{}_1}{2}} d\chi^2{}_{\phi-1}$$

$$= A_{\phi-1}A_1 \int_0^{\chi^2\phi} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-1}{2}-1} e^{\frac{-\chi^2{}_{\phi-1}}{2}} \cdot A_1 (\chi^2{}_1)^{\frac{1}{2}-1} e^{\frac{-\chi^2{}_1}{2}} d\chi^2{}_{\phi-1}$$

$$= A_{\phi-1}A_1 \int_0^{\chi^2\phi} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-3}{2}} e^{\frac{-\chi^2{}_{\phi-1}}{2}} \cdot \left(\chi^2{}_\phi - \chi^2{}_{\phi-1}\right)^{\frac{-1}{2}} e^{\frac{-\left(\chi^2{}_\phi - \chi^2{}_{\phi-1}\right)}{2}} d\chi^2{}_{\phi-1}$$

$$= A_{\phi-1}A_1 \int_0^{\chi^2\phi} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-3}{2}} \left(\chi^2{}_\phi - \chi^2{}_{\phi-1}\right)^{\frac{1}{2}-1} e^{\frac{-\chi^2{}_{\phi-1} - \left(\chi^2{}_\phi - \chi^2{}_{\phi-1}\right)}{2}} d\chi^2{}_{\phi-1}$$

$$= A_{\phi-1}A_1 \int_0^{\chi^2\phi} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-3}{2}} \left(\chi^2{}_\phi - \chi^2{}_{\phi-1}\right)^{\frac{-1}{2}} e^{\frac{-\chi^2{}_\phi}{2}} d\chi^2{}_{\phi-1}$$

$$= A_{\phi-1}A_1 e^{\frac{-\chi^2{}_\phi}{2}} \int_0^{\chi^2\phi} \left(\chi^2{}_{\phi-1}\right)^{\frac{\phi-3}{2}} \left(\chi^2{}_\phi - \chi^2{}_{\phi-1}\right)^{\frac{-1}{2}} d\chi^2{}_{\phi-1}$$

Here,

$$\chi^2{}_{\phi-1} = u\chi^2{}_{\phi}$$

$$d\chi^2{}_{\phi-1} = \chi^2{}_{\phi} du$$

$$\mathrm{P}\left(\chi^2{}_{\phi}\right) = A_{\phi-1} A_1 e^{\frac{-\chi^2{}_{\phi}}{2}} \int_0^1 \left(u\chi^2{}_{\phi}\right)^{\frac{\phi-3}{2}} \left(\chi^2{}_{\phi} - u\chi^2{}_{\phi}\right)^{\frac{-1}{2}} \chi^2{}_{\phi} du$$

$$= A_{\phi-1} A_1 e^{\frac{-\chi^2{}_{\phi}}{2}} \int_0^1 u^{\frac{n-3}{2}} \chi^2{}_{\phi}{}^{\frac{n-3}{2}} \chi^2{}_{\phi}{}^{\frac{-1}{2}} (1-u)^{\frac{-1}{2}} \chi^2{}_{\phi} du$$

$$= A_{\phi-1} A_1 e^{\frac{-\chi^2{}_{\phi}}{2}} \int_0^1 \chi^2{}_{\phi}{}^{\frac{\phi}{2}-1} u^{\frac{\phi-3}{2}} (1-u)^{\frac{-1}{2}} du$$

$$= A_{\phi-1} A_1 \chi^2{}_{\phi}{}^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_{\phi}}{2}} \int_0^1 u^{\frac{n-3}{2}} (1-u)^{\frac{-1}{2}} du$$

$$\mathrm{P}(\chi^2{}_{n}) = A_{\phi} \left(\chi^2{}_{\phi}\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_{\phi}}{2}} = A_{\phi-1} A_1 \chi^2{}_{\phi}{}^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_{\phi}}{2}} \int_0^1 u^{\frac{n-3}{2}} (1-u)^{\frac{-1}{2}} du$$

$$A_{\phi} = A_{\phi-1} A_1 \int_0^1 u^{\frac{n-3}{2}} (1-u)^{\frac{-1}{2}} du$$

$$A_1 = \frac{1}{\sqrt{2\pi}}$$

$$A_{\phi} = A_{\phi-1} \frac{1}{\sqrt{2\pi}} \int_0^1 u^{\frac{n-3}{2}} (1-u)^{\frac{-1}{2}} du$$

From this we can express relation between A_{ϕ} and ϕA_{-1}, and we can express the probability as

$$\mathrm{P}(\chi^2{}_{\phi}) = A_{\phi} \chi^2{}_{n}{}^{\frac{n}{2}-1} e^{\frac{-\chi^2{}_{n}}{2}}$$

However, our purpose is to obtain general expression of $\mathrm{P}(\chi^2{}_{\phi})$, and we should express A_{ϕ} only by $\chi^2{}_{\phi}$ and n.

$\mathrm{P}(\chi^2{}_{\phi})$is probability and $0 \leq \mathrm{P}(\chi^2{}_{n}) \leq 1$. So, the infinite integral is 1

$$\int_0^{\infty} \mathrm{P}(\chi^2{}_{\phi})\, \mathrm{dz} = 1$$

$$\int_0^{\infty} A_{\phi} \left(\chi^2{}_{\phi}\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_{\phi}}{2}} d\chi^2{}_{\phi} = 1$$

$$A_{\phi} = \frac{1}{\int_0^{\infty} \left(\chi^2{}_{\phi}\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_{\phi}}{2}} d\chi^2{}_{\phi}}$$

Denominator of the formula is gamma function. We can express chai square distribution simply using gamma function.

Definition of gamma function is as follow

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

$$\text{z: complex number, z} \geq 0 \text{ on complex plane}$$

Formula18

Γ function was formulated by Leohard Euler. Followings are commonly used nature of Γ function.

$$\Gamma(n+1) = n!$$

$$\Gamma(1) = 1$$

$$\Gamma\left(\frac{1}{2}\right) = \pi$$

$$\Gamma\left(\frac{1}{2} + n\right) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}$$

Formula 19

We consider the formula of chai square using Γ function

$$A_\phi = \frac{1}{\int_0^\infty \left(\chi^2{}_\phi\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}} d\chi^2{}_\phi}$$

about

$$\int_0^\infty \left(\chi^2{}_\phi\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}} \mathrm{d}\chi^2{}_\phi$$

Here, $\chi^2{}_\phi = 2\mathrm{t}$

$$\mathrm{d}\chi^2{}_\phi = 2\mathrm{dt}$$

$$\int_0^\infty (2t)^{\frac{\phi}{2}-1} e^{-t} 2\mathrm{dt}$$

$$= \int_0^\infty 2^{\frac{\phi}{2}-1} 2t^{\frac{\phi}{2}-1} e^{-t} \mathrm{dt}$$

$$= 2^{\frac{\phi}{2}} \int_0^\infty t^{\frac{\phi}{2}-1} e^{-t} \mathrm{dt}$$

$$\Gamma\left(\frac{\phi}{2}\right) = \int_0^\infty t^{\frac{\phi}{2}-1} e^{t} dt$$

$$A_\phi = \frac{1}{\int_0^\infty (z)^{\frac{\phi}{2}-1} e^{\frac{-z}{2}} dz}$$

$$A_\phi = \frac{1}{2^{\frac{\phi}{2}} \Gamma\left(\frac{\phi}{2}\right)}$$

$$\int_0^\infty \left(\chi^2{}_n\right)^{\frac{\phi}{2}-1} e^{\frac{-\chi^2{}_\phi}{2}} \mathrm{d} = 2^{\frac{\phi}{2}} \Gamma\left(\frac{\phi}{2}\right)$$

$$A_\phi = \frac{1}{\int_0^\infty (z)^{\frac{\phi}{2}-1} e^{\frac{-z}{2}} dz}$$

$$A_\phi = \frac{1}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}$$

$$\mathrm{P}({\chi^2}_\phi) = \frac{{{\chi^2}_\phi}^{\frac{\phi}{2}-1}}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} e^{\frac{-{\chi^2}_\phi}{2}}$$

Formula 20

Chai square distribution is used in chai square test to deny the correct ness of expectation value obtained from a null hypothesis (to deny null hypothesis). In chai square test, definition of chai square is as follow.

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i}$$

e_i: expectation value of factor i

Formula 21

Theoretical definition obtained from normal distribution is as follow.

$$\chi^2 = \sum \frac{(f_i - \mu_i)^2}{{\sigma_i}^2}$$

We have to discuss similarity of two definition of χ^2.

We use following simplified calculation of variance

$$V_x = \mathrm{E}(x^2) - \mathrm{E}(x)^2$$

V_x: variance of x

Formula 14

The proof of the simplified calculation is as follow.

Definition of V_x

$$\begin{aligned} V_x &= \mathrm{E}\big(x - \mathrm{E}(x)\big)^2 \\ &= \mathrm{E}(x^2) - \mathrm{E}\big(2x\mathrm{E}(x)\big) + \mathrm{E}(x)^2 \\ &= \mathrm{E}(x^2) - 2\mathrm{E}(x)\mathrm{E}(x) + \mathrm{E}(x)^2 \\ &= \mathrm{E}(x^2) - \mathrm{E}(x)^2 \end{aligned}$$

Expectation value is sum of products of value and probability. Expectation value of z^k is as follow.

$$\mathrm{E}(z^k) = \int_0^\infty z^k P(z) dz$$

In the case of chai square distribution

$$\mathrm{E}(z^k) = \int_0^\infty z^k \frac{z^{\frac{\phi}{2}-1}}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} e^{\frac{-z}{2}} dz$$

$$= \frac{1}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} \int_0^\infty z^k z^{\frac{n}{2}-1} e^{\frac{-z}{2}} dz$$

$$= \frac{1}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} \int_0^\infty z^{\frac{\phi}{2}+k-1} e^{\frac{-z}{2}} dz$$

$$= \frac{1}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} \int_0^\infty (2t)^{\frac{\phi}{2}+k-1} e^{-t} 2dt$$

$$\because \ \frac{z}{2} = t、\frac{dz}{dt} = 2$$

$$= \frac{1}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} \int_0^\infty 2^{\frac{\phi}{2}+k-1} 2\, t^{\frac{\phi}{2}+k-1} e^{-t} dt$$

$$= \frac{1}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} \int_0^\infty 2^{\frac{\phi}{2}+k}\, t^{\frac{\phi}{2}+k-1} e^{-t} dt$$

$$= \frac{2^{\frac{\phi}{2}+k}}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} \int_0^\infty t^{\frac{\phi}{2}+k-1} e^{-t} dt$$

Here,

$$\mathrm{E}(z^k) = \frac{2^{\frac{\phi}{2}+k}\Gamma\left(\frac{\phi}{2}+k\right)}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}$$

$$\mathrm{E}(z^0) = \frac{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)} = 1$$

$$\mathrm{E}(z^1) = \frac{2^{\frac{\phi}{2}+1}\Gamma\left(\frac{\phi}{2}+1\right)}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}$$

$$= \frac{2\frac{\phi}{2}\Gamma\left(\frac{\phi}{2}\right)}{\Gamma\left(\frac{\phi}{2}\right)}$$

$$\because \Gamma(z+1) = z\Gamma(z)$$

$$= \phi$$

$$E(z^2) = \frac{2^{\frac{\phi}{2}+2}\Gamma\left(\frac{\phi}{2}+2\right)}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}$$

$$= \frac{2^{\frac{\phi}{2}+2}\left(\frac{\phi}{2}+1\right)\Gamma\left(\frac{\phi}{2}+1\right)}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}$$

$$= \frac{2^{\frac{\phi}{2}+2}\left(\frac{\phi}{2}+1\right)\left(\frac{\phi}{2}\right)\Gamma\left(\frac{\phi}{2}\right)}{2^{\frac{\phi}{2}}\Gamma\left(\frac{\phi}{2}\right)}$$

$$= \frac{2^{2}\left(\frac{\phi}{2}+1\right)\left(\frac{\phi}{2}\right)\Gamma\left(\frac{\phi}{2}\right)}{\Gamma\left(\frac{\phi}{2}\right)}$$

$$= \phi(\phi+2)$$

From $E(z^2)$ *and* $E(z)^2$, we can obtain variance by following simplified calculation.

$$V(z) = E(z^2) - E(z)^2$$

$$= \phi(\phi+2) - \phi^2$$

$$= 2\phi$$

Conclusively, mean and variance is as follows

$$\mu = \phi$$

$$\sigma^2 = 2\phi$$

When we presume $e_i = \mu_i$

$$e_i = \mu_i = \phi = \frac{\sigma^2}{2}$$

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i} = \sum \frac{2(f_i - \mu_i)^2}{{\sigma_i}^2}$$

This is similar to the definition of chai square excepting existence of 2 in numerator. In the process of making chai square distribution, we fold normal distribution at expectation value, though $(f_i - e_i)^2$ is variance of one side, the same variance exists in the opposite side of the expectation value. For this reason, we have to multiply 2 to $(f_i - e_i)^2$ to normalize the variance. We can conclude that $\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i}$ which obtained from observed data as estimated value of $\chi^2 = \sum \frac{(f_i - \mu_i)^2}{\sigma_i^2}$ which is theoretical definition of χ^2.

Comments from author

Many text books lack explanation of the relation between theoretical definition of χ^2 and χ^2 in chai square test. so readers cannot understand the meaning of chai square analysis. This is because, we need to explain Γ function in the process to derive chai square distribution from normal distribution. Γ function is a transcendental function and very difficult to explain. When we accept e_i only as mean of the data, we cannot understand the meaning of $\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i}$. The author is thinking that it is better to explain the process to derive chai square distribution from normal distribution for accurate understanding of chai square test.