

III-2-5. カイ二乗分布

正規分布は以下のように書けます。

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

この式で、 $\frac{x-\mu}{\sigma}$ の意味は標準化で、標準正規分布を作っています。その結果、すべて標準正規分布の確率分布で論ずることが出来るようになります。カイ二乗分布はこの標準化された 2 乗値を一つの変数として、その確率分布を論じたものです。

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

この標準化した距離を x とするならば

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

となるのだから、新しい変数 $\left(\frac{x-\mu}{\sigma}\right)^2$ は x^2 (エックススクエア) ではないかと考える人もいるでしょう。 χ^2 と新たな変数として表現することによって、この二乗値を一つの変数とするということを明瞭にしているのだと思います。つまり、 $P(x)$ が正規分布に従うときに $P(\chi^2)$ はカイ二乗分布に従うということです。

普通 χ^2 検定は、期待値には程遠いという検定に使われます。 χ^2 の値がある値以上である場合に、得られた値が期待値から統計的に遠いという判定をします。つまり、同じ集団から統計的にそんな大きな値が得られることは統計的にないと結論するのです。もっと、わかりやすく言えば、そのような期待は間違っていると結論します。最も重要なカイ二乗分布の性質は、再生性で、二つの変数の合計が、元の変数が従っていた分布に従うということです。

$$X_1 \sim \chi^2_m, X_2 \sim \chi^2_n \rightarrow X_1 + X_2 \sim \chi^2_{m+n}$$

それでは、 χ^2_1 の公式を作ります。

χ^2_1 の意味は、自由度 1 の χ^2 分布ということです。そうである場合とそうでない場合の 2 つの二つの変数があるのだけれど、一つの変数はもう一方の余事象だから、全体の大きさが決まれば、変動する変数は 1 つだけということになるので、自由度 1 ということになります。これはそのまま正規分布を書き換えればよいのですが、

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

$x = 0$ から $x = x$ までの累積確率は

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

になるのですが $x^2 = (-x)^2$ ですから、反対側にも同じ大きさの確率分布があって、それも合わせて考えなくてはならないから、0 を中心に折り返して、重ね合わせて、さらに 2 乗だから、2 乗分になるように横方向に引っ張った形になります。

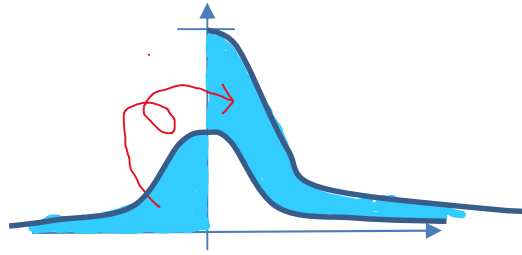


図 18. 左右の確率の重ね合わせ

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

確率は累積確率の微分だから

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

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

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

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

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

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

これで、自由度 1 の χ^2 分布出来上がります。

自由度 2 の場合を考えます。

ある現象 x とある現象 y があって、どちらも正規分布に従う場合のことです。

$x \sim N(0,1)$ and $y \sim N(0,1)$.

互いに独立なのだから、それが同時の起こる確率は掛け算になります。s.

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

混乱しないように一つは χ^2_x 、もう一つは χ^2_y と書くことにして、自由度 1 を強調して、 f $\chi^2_{1:x}$ 、 $\chi^2_{1:y}$ と書くことになり、求める確率分布の変数は $\chi^2_{2:x+y}$ となります。

これを使って $P(\chi^2_{2:x+y})$ を公式化します。

無限積分で書くと.

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

となりますが、このまま右辺を積分するのは難しいので、積分方向が $\chi^2_{1:x} + \chi^2_{1:y}$ と並行線とそれに直行するベクトルの方向になるように座標軸を回転することを考えます。そのような回転は、 $-\frac{\pi}{4}$ の回転です。これなら重積分できそうです。

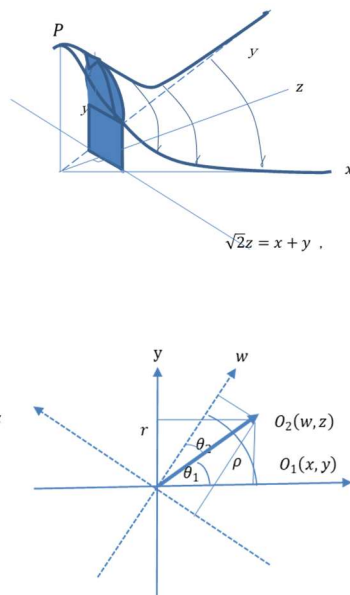


図 19. 回転による座標軸の変換

つまり $O_1(x, y)$, という座標系を $P_1(r, \theta_1)$ という極座標変換して回転し、その結果得られた $P_2(r, \theta_2)$ を直交座標に戻すということです。

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

$$x = r \cos \theta_1$$

$$y = r \sin \theta_1$$

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

$$w = r \cos \theta_2$$

$$z = r \sin \theta_2$$

$$\theta_2 = \theta_1 - \rho$$

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

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

ここで $\rho = -\frac{\pi}{4}$ ですから

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

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

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

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

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

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

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

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

無限積分を考えれば

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

となりますが、ここでは不定積分 $\int \int P(\chi^2_{1:x})P(\chi^2_{1:y})d\chi^2_{1:x}d\chi^2_{1:y}$ について考えます。

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

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

$$\begin{aligned} & \int \int P(\chi^2_{1:x})P(\chi^2_{1:y})d\chi^2_{1:x}d\chi^2_{1:y} \\ &= \int \int \frac{1}{\sqrt{2\pi\chi^2_{1:x}}} e^{-\frac{\chi^2_{1:x}}{2}} \cdot \frac{1}{\sqrt{2\pi\chi^2_{1:y}}} e^{-\frac{\chi^2_{1:y}}{2}} d\chi^2_{1:x}d\chi^2_{1:y} \\ &= \frac{1}{2\pi} \int \int \frac{1}{\sqrt{\chi^2_{1:x}\chi^2_{1:y}}} e^{-\frac{(\chi^2_{1:x}+\chi^2_{1:y})}{2}} d\chi^2_{1:x}d\chi^2_{1:y} \\ &= \frac{1}{2\pi} \int \int \frac{1}{\sqrt{\chi^2_{1:x}\chi^2_{1:y}}} e^{-\frac{\chi^2_{2:x+y}}{2}} d\chi^2_{1:x}d\chi^2_{1:y} \end{aligned}$$

我々の「目的は z のベクトル上で 0 から Z の範囲で積分し、それを、 z で微分するという

ことです。それは、 $y = \sqrt{2}z - x$ 上で、 $w-z$ 平面に直交する断面の断面積を $-z$ から $+z$ の範囲で積分して、それを z のベクトル方向に積分すれば得られるでしょう。

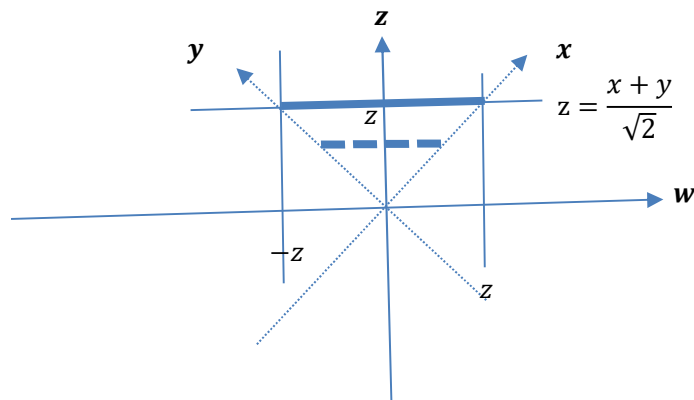


図 20. $w-z$ 平面上の積分方向

変形

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

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

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

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

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

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

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

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

積分領域を与えます。

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

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

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

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

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

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

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

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

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

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

結論として

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

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

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

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

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

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

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

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

となります。これで

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

ということは確かめられたのですが

0 から $\chi^2_{2:x+y}$ までの累積確率を求めると

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

これを微分すれば確率が得られます。

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

これで、 $P(\chi^2_2)$ の公式が得られました。これを繰り返せば、どのような自由度のカイ二

乗分布を作れますが、少し時間がかかりすぎます。 $P(\chi^2_\phi)$ の一般的な公式を作ること

をえます。 ϕ は自由度です。自由度はすでに説明したように、余事象を考えた時に変動する変数の数です。ボールがいくつかあって、ボールの色が赤白2種類で、赤白のボールの和を数が出る頻度考えるならば自由度1ですが、赤白緑のボールがあって、それぞれの色のボールが出る頻度を考えるならば自由度2です。

$\phi = 1$ の時

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

$\phi = 2$ の時

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

$\phi = 3$ の時

$\phi = 2$ の時にやったことを繰り返します。

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

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

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

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

$$\begin{aligned} & \int \int P(\chi^2_1) P(\chi^2_2) d\chi^2_2 d\chi^2_1 \\ &= \int \int \frac{1}{\sqrt{2\pi\chi^2_1}} e^{-\frac{\chi^2_1}{2}} \cdot \frac{1}{2} e^{-\frac{\chi^2_2}{2}} d\chi^2_2 d\chi^2_1 \end{aligned}$$

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

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

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

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

ここで, $\chi^2_2 = u\chi^2_3$ ですから

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

$$0 \leq u \leq 1$$

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

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

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

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

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

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

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

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

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

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

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

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

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

$P(\chi^2_1), P(\chi^2_2), P(\chi^2_3)$ を比較します

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

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

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

一般化して書きたいので、法則を探します

$$P(\chi^2_n) = A_n(\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}}$$

と書けることにします。すると

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

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

$$P(\chi^2_n) = A_n(\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}}$$

ここで χ^2_{n-1} は 0 から χ^2_n の領域で変動しまから

$$P(\chi^2_n) = \int_0^{\chi^2_n} P(\chi^2_1) P(\chi^2_{n-1}) d\chi^2_{n-1}$$

$$A_n(\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}} = \int_0^{\chi^2_n} A_{n-1}(\chi^2_{n-1})^{\frac{n-1}{2}-1} e^{-\frac{\chi^2_{n-1}}{2}} \cdot A_1(\chi^2_1)^{\frac{1}{2}-1} e^{-\frac{\chi^2_1}{2}} d\chi^2_{n-1}$$

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

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

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

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

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

ここで次のような変換を考えます

$$\begin{aligned}\chi^2_{n-1} &= u\chi^2_n \\ d\chi^2_{n-1} &= \chi^2_n du\end{aligned}$$

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

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

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

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

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

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

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

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

これで、 $P(\chi^2_n) = A_n \chi^2_n^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}}$ と書けることがわかりました。しかしもっと一般的な記述

にしないと、なんだかよくわかりません。

χ^2_n と n だけで書き表すということです。 .

$P(\chi^2_n)$ は確率でその無限定積分は1になるはずだから

$$\int_0^\infty P(\chi^2_n) dz = 1$$

$$\int_0^\infty A_n (\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}} d\chi^2_n = 1$$

$$A_n = \frac{1}{\int_0^\infty (\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}} d\chi^2_n}$$

この分母はガンマ関数と言われているものです。これを使えばもう少し簡単に書けます。

以下はガンマ関数の定義です

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

z : は複素数で複素平面上で正

式 18

Γ 関数は大数学者オイラーが考えた関数です。以下はガンマ関数の代表的な性質です。

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

$$\Gamma(1) = 1$$

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

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

式 19

ガンマ関数を使ってカイ二乗分布を書きます。

$$A_n = \frac{1}{\int_0^{\infty} (\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}} d\chi^2_n}$$

$\int_0^{\infty} (\chi^2_n)^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}} d\chi^2_n$ について

$$\chi^2_n = 2t$$

$$d\chi^2_n = 2dt$$

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

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

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

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

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

$$A_n = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}$$

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

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

$$A_n = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}$$

$$P(\chi^2_n) = \frac{\chi^2_n^{\frac{n}{2}-1} e^{-\frac{\chi^2_n}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}$$

式 20

データから実際にカイ二乗値を計算するときは、

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

f :観測値 e :期待値

式 21

分子の方はカイ二乗値の定義から十分納得できますが、分母が何故期待値なのかわかりません。これについて考えます。統計分析の実務としては、こんなことを考える必要はありませんが、統計的な確率とは何かを理解する上では、この考察は重要です。しっかり理解してください。

期待値というのは値と確率の積の総和ですから、

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

カイ二乗の場合は

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

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

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

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

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

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

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

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

ここで

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

だから

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

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

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

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

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

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

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

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

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

$$= n(n+2)$$

分散を簡易計算の方法で求めると、

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

$$\begin{aligned} &= n(n+2) - n^2 \\ &= 2n \end{aligned}$$

結論としては

$$\begin{aligned} \mu &= n \\ \sigma^2 &= 2n \end{aligned}$$

μ は期待値だから

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

つまり、分母は期待値の意味は分散の2分の1なのです。だから、カニ2乗はF値と同じような分散比なのです。