Hypothesis testing about vector mean

نعلم انه في حالة التوزيع الطبيعي الاحادي المتغير عندما يتم اختيار فرضية معينة حول متوسط المجتمع الطبيعي ويكون تباينه معلوماً وفترة ثقة لذلك المتوسط. سوف تحتوي على الفرق بين متوسط العينة ومتوسط المجتمع له توزيع طبيعي قياسي لذلك يتم الاعتماد على جداول التوزيع الطبيعي القياسي من خلال معرفة مستوى المعنوية وبناء عليه يتم اجراء الاختبار وتكوين فترة الثقة.

وكذلك التوزيع الطبيعي المتعدد المتغيرات يتم استخدام فرق بين متوسط العينة ومتوسط المجتمع ($\neq$) وتكون معلومة هذا الفرق يتبع توزيع طبيعي متعدد متغيرات معلوم

وبإمكاننا وضع حدود يتم على ضوئها قبول او رفض الفرضية الاحصائية، وتعتمد تلك الحدود على التوزيع الاحتمالي الشكل التربيعي لذلك التوزيع وذلك الشكل يتبع توزيع وبع كاي بدرجات حرية مساوية لعدد متغيرات المجتمع العشوائي ويتم قبول او رفض الفرضية بنفس الاسلوب المتبع في حالة المتغير المتعدد. وفيما يتعلق بفترة الثقة سوف يكون لدينا حالات وفترات الثقة وسوف يتم اختيار الثقة سوف يكون لدينا حالات وفترات الثقة وسوف يتم اختيار فترة ثقة واحدة تحتوي متباينة معينة سوف يتم توضيح لاحقاً.

Scanned with CamScanner

وكما نعلم في حالة المتغير المعتاد تكون فترات الثقة واختبار الفرضيات وعند تكوين فترة الثقة حول وسط المجتمع الصفري عندما يكون التباين معلوم نستخدم ثقة $(1-\alpha)$ اذا ان (L, U) فيكون بالشكل الاتي:

$$(L, U) = \bar{X} \mp Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

وعند اختبار الفرضية:

$$H_0 : \mu = \mu_0$$

كانت الاختبار المستخدم هو:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

هذا الاختبار يتبع التوزيع الصفري القياسي

وعلى هذا الاساس نقارن قيمة المتغير الاحصائي Z مع القيمة الجدولية عند مستوى معنوية $\alpha/2$

ممكن تعميم حالة الاختبار اعلاه اي حالة تعدد المتغيرات اذ ان يمكن تعميم المتغير الاحصائي اللائم لاختبار الفرضية H_0 على هيئة $(\mu = \mu_0)$

له شكل اخر وسوف يتم استخدامه بالاعتماد على النظرية الاتية

Theorem:

If $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, then $(\underline{X} - \underline{\mu})' \Sigma^{-1} (\underline{X} - \underline{\mu})$

$$\sim \chi^2_{(p)}$$

اذا كانت جميع المتغيرات العشوائية تتبع توزيع طبيعي فان الشكل التربيعي $(\underline{X} - \underline{\mu})' \Sigma^{-1} (\underline{X} - \underline{\mu})$ سوف يتبع توزيع [illegible] بدرجات حرية (p)

Scanned with CamScanner

Proof :

by using m.g.f ..

$$f(\underline{x}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})}$$

Let $Q = (\underline{x}-\underline{\mu}) \Sigma^{-1} (\underline{x}-\underline{\mu})$

$$M_{\phi}(t) = E e^{tQ} = \int_{-\infty}^{\infty} e^{tQ} f(x) dx$$

$$= \int_{-\infty}^{\infty} \frac{e^{t(\underline{x}-\underline{\mu}) \Sigma^{-1} (\underline{x}-\underline{\mu})}}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}) \Sigma^{-1} (\underline{x}-\underline{\mu})} dx$$

عند الجمع نجمع الاسس واقرأ عامل مشترك كما يلي

$$= \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{p/2} |\Sigma|^{\frac{1}{2}}} e^{(t-\frac{1}{2})[\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})]} dx$$

$$t - \frac{1}{2} = \frac{2t-1}{2} = \frac{-(1-2t)}{2}$$

$$= \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{p/2} |\Sigma|^{\frac{1}{2}}} e^{\frac{-(1-2t)}{2} (\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})} dx$$

---(1)

Scanned with CamScanner

$$= \frac{1}{(2\pi)^{p/2} |\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{\infty} \underbrace{e^{-\frac{(1-2t)}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}}_{\text{Kernal نواة}} dx$$

---(1)

التكامل في المعادلة اعلاه يمثل تكامل نواة التوزيع الطبيعي المتعدد حيث

μ وتباين $\left(\frac{\Sigma}{1-2t}\right)$

* ملاحظة : النواة = الثابت $\times$ مقلوب الثابت بعد التغير وذلك

بالاستعانة بالمتباينة $\left(\frac{1}{1-2t}\right)^{\frac{1}{2}}$ وهو $\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)$

$$\therefore M_Q(t) = \underbrace{\frac{1}{(2\pi)^{p/2}|\Sigma|^{\frac{1}{2}}}}_{\text{الثابت}} \cdot \underbrace{(2\pi)^{p/2}\left[\frac{|\Sigma|}{|1-2t|}\right]^{\frac{1}{2}}}_{\text{مقلوب الثابت بعد التغير}}$$

وباستخدام اهم خصائص المحدد وهي اذا كانت المصفوفة A مصفوفة

ربعية ذات حجم $(n \times n)$ وكان لدينا الثابت c فان

$$|cA| = c^n |A|$$

فلو : $A_{3\times 3}$ اذن

$$|cA_{3\times 3}| = c^3 |A|$$

وعليه يكون دالة المولد الكامنة

$$M_a(t) = \frac{1}{|\Sigma|^{\frac{1}{2}}}\left[\frac{|\Sigma|}{(1-2t)^p}\right]^{\frac{1}{2}} = (1-2t)^{-p/2}$$

Scanned with CamScanner

which is the m.g.f of χ^2 distribution with (P) d.f (تقريبا)

$$\therefore (\underline{X}-\underline{\mu})' \underline{\Sigma}^{-1} (\underline{X}-\underline{\mu}) \sim \chi^2_{(P)}$$

Example

A r.s. of size n taken from $N_p(\underline{\mu}, \underline{\Sigma})$ and if $\bar{\underline{x}}$ be the estimator for $\underline{\mu}$. Find the P.d.f of $n(\bar{\underline{x}}-\underline{\mu})' \underline{\Sigma}^{-1} (\bar{\underline{x}}-\underline{\mu})$

Sol

$$\bar{\underline{x}} = \frac{\sum_{i=1}^{n} \underline{x}_i}{n} = \sum_{i=1}^{n} \frac{\underline{x}_i}{n}$$

we know that $\underline{x}_i \sim N_p(\underline{\mu}, \underline{\Sigma})$ for $i = 1, 2, \ldots, n$

$$\frac{\underline{x}_i}{n} \sim N_p\left(\frac{\underline{\mu}}{n}, \frac{\underline{\Sigma}}{n}\right)$$

$$f(\bar{\underline{x}}) = \frac{1}{(2\pi)^{p/2} \left|\frac{\underline{\Sigma}}{n}\right|^{1/2}} e^{-\frac{n}{2}(\bar{\underline{x}}-\underline{\mu})' \underline{\Sigma}^{-1} (\bar{\underline{x}}-\underline{\mu})}$$

بعد تبسيط المقدار الثابت نجد أن

$$f(\bar{\underline{x}}) = \frac{1}{(2\pi)^{p/2} \left(\frac{1}{n}\right)^{p/2} |\underline{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\bar{\underline{x}}-\underline{\mu})' \underline{\Sigma}^{-1} (\bar{\underline{x}}-\underline{\mu})}$$

Scanned with CamScanner

Let $Q = -n(\bar{X}-\underline{\mu})' \Sigma^{-1} (\bar{X}-\underline{\mu})$

$$M_Q(t) = E e^{tQ} = \int_{-\infty}^{\infty} e^{tQ} f(\underline{x})\, d\bar{x}$$

$$= \frac{(n)^{p/2}}{(2\pi)^{p/2} |\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{\infty} e^{n(\bar{X}-\mu)' \Sigma^{-1} (\underline{x}-\underline{\mu})} e^{-\frac{n}{2}(\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})} d\bar{x}$$

$$= C \int_{-\infty}^{\infty} e^{-\frac{(1-2t)}{2} n(\bar{X}-\underline{\mu})' \Sigma^{-1} (\bar{X}-\underline{\mu})} d\bar{x}$$

دالة التوزيع الطبيعي المتعدد (kernal) نواة التكامل عبارة عن داخل التكامل

$$\left(\frac{\Sigma}{n(1-2t)}\right)$$ وهو يكافئ Σ عينة حجم

$$\therefore M_Q(t) = C(2\pi)^{p/2} \left|\frac{\Sigma}{n(1-2t)}\right|^{\frac{1}{2}}$$

وحسب خاصية المحددات نجد أن

$$M_Q(t) = C(2\pi)^{p/2} |\Sigma|^{\frac{1}{2}} [n(1-2t)]^{-p/2}$$

$$\therefore M_Q(t) = \frac{(n)^{p/2} (2\pi)^{p/2} |\Sigma|^{1/2} (n)^{-p/2} (1-2t)^{-p/2}}{(2\pi)^{p/2} |\Sigma|^{1/2}} = (1-2t)^{-p/2}$$

Scanned with CamScanner

which is the m.g.f of $\chi^2_{(p)}$

$$\therefore f(Q) = \frac{\left(\frac{1}{2}\right)^{p/2}}{\Gamma\frac{p}{2}} Q^{\frac{p}{2}-1} e^{-\frac{Q}{2}}$$

$Q > 0$

p is Positive integer

Hypothesis tests about the vector mean when Σ is known

اذا كان لدينا عينة عشوائية ذات حجم (n) والمجتمع طبيعي حيث متجه متوسط μ ومصفوفة تباين Σ وان مقدر الامكانة الاعظم لمتجه المتوسطات هو:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

ولعرض اختبار الفرضية

$$H_0 : \mu = \mu_0$$

vs $H_1 : \mu \neq \mu_0$ when C (Σ known)

باستخدام نسبة الامكانة التي تمثل الامكانة عند فضاء المعلم H_0 كما والف الامكانة عند الفرضية البديلة H_1 اقل او تساوي الثابت C

$$\frac{\text{دالة الامكانة تحت } H_0}{\text{دالة الامكانة تحت } H_1} = \frac{L_{max}(\mu = \mu_0, \Sigma)}{L_{max}(\mu \neq \mu_0, \Sigma)} \leq C$$

Scanned with CamScanner

باستخدام إحصائية الاختبار كمثل لاختبار الفرضيات اللازم لاختبار الفرضية العادية (تكون دائماً فيه مولف)

$$W = n(\bar{X}-\underline{\mu}_0)' \Sigma^{-1} (\bar{X}-\underline{\mu}_0) \sim \chi^2_{(p)}$$

حيث قارن قيمة الاختبار المحسوبة (W) [فيه مولف دائماً اكبر من الصفر] بقيمة χ^2 الجدولية ذات درجة حرية (p) وبمستوى معنوية (α) فاذا كانت قيمة (W) أكبر من الجدولية تكون القرار يقبل (H_1) وبعكسه نرفض (H_0) ونقبل (H_1)

Example

A rvs of size (25) taken from $N_2(\underline{\mu}, \begin{bmatrix} 16 & 8 \\ & 9 \end{bmatrix})$ and the sample vector mean $\bar{X} = [15.4 \quad 9.9]'$ test the hypothesis at $\alpha = 0.05$: $H_0: \underline{\mu} = \begin{bmatrix} 17 \\ 10 \end{bmatrix}$ V.S $H_1: \underline{\mu} \neq \begin{bmatrix} 17 \\ 10 \end{bmatrix}$

Sol

$$W = n(\bar{X}-M)' \Sigma^{-1} (\bar{X}-M)$$

$$|\Sigma| = (16)(9) - (8)(8) = 80$$

$$\therefore \Sigma^{-1} = \begin{bmatrix} \frac{9}{80} & -\frac{8}{80} \\ -\frac{8}{80} & \frac{16}{80} \end{bmatrix}$$

$$\therefore W = \frac{25}{80} \begin{bmatrix} 15.4 \\ 9.9 \end{bmatrix} - \begin{bmatrix} 17 \\ 10 \end{bmatrix}' \begin{bmatrix} 9 & -8 \\ -8 & 16 \end{bmatrix} \begin{bmatrix} \begin{pmatrix} 15.4 \\ 9.9 \end{pmatrix} - \begin{pmatrix} 17 \\ 10 \end{pmatrix} \end{bmatrix}$$

Scanned with CamScanner

$W = 6.45$

نقارن W مع القيمة الجدولية $\chi^2_{(2,0.05)} = 5.991$

$\therefore W > 5.991$ Thus we reject H_0

$\therefore$ نرفض فرضية العلم اي ان $\mu \neq \begin{bmatrix} 7 \\ 10 \end{bmatrix}$

ملاحظة

ملاحظة هل ان (μ^*) تقع ضمن منطقة الثقة :-

اعرف لدينا المتجه (μ^*) عبارة عن مجموعة من المتجهات التي تحقق المتباينة التالية :-

$$n(\bar{x} - \mu^*)' \Sigma^{-1} (\bar{x} - \mu^*) \leq \chi^2_{p(\alpha)}$$

μ^* تقول $\chi^2_{p(\alpha)}$ اذا كانت العلاقة اقل من القيمة الجدولية

Example

بتقديرنا اذا

Check wether $(9 \quad 9)$ lies in the Confidence rejon or (interval) or not.

Solu

$$\frac{25}{80} \begin{bmatrix} 15.4 - 9 \\ 9.9 - 9 \end{bmatrix}' \begin{pmatrix} 9 & -8 \\ -8 & 16 \end{pmatrix} \begin{pmatrix} 15.4 - 9 \\ 9.9 - 9 \end{pmatrix} \leq \chi^2_{(2,0.05)}$$

$$90.45 > 5.993$$

Thus $\mu^* = \begin{bmatrix} 9 \\ 9 \end{bmatrix}$ lies outside the Confidence interval

Scanned with CamScanner

2- Test of hypothesis about the vector when (Σ is uknown)

نقوم في حالة التوزيع الطبيعي الاحادي المتغير وعندما تكون σ غير معلومة واختبار الفرضية $H_0: M = M_0$ ستقوم الحيز الاحصائي

$$t = \frac{\bar{X} - M_0}{\frac{S}{\sqrt{n}}}$$

هذا الحيز يتبع توزيع t بدرجة حرية $(n-1)$ ويتمثل القرار حول قبول او رفض الفرضية عبارة عن مقارنة قيمة الحيز مع القيمة الجدولية.

في حالة المتجه المتغيرات واختبار الفرضية $H_0: M = M_0$ ضد الفرضية البديلة $H_1: M \neq M_0$ الاختبار الذي سنتبع اقترح العالم (Hotelling) الاحصائية تعرف بالشكل الاتي:

$$T^2 = n(\bar{X} - M_0)' S^{-1} (\bar{X} - M_0)$$

وتدعى الاحصائية (T^2 Hotelling T test)

اي ان Σ غير معلوم الحيز لـ Σ اي:

اذا ان المصفوفة S تمثل المتجه غير

$$S = \frac{A}{n-1} \quad , \quad A = \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})'$$

$$= XX' - n\bar{X}\bar{X}' = \begin{bmatrix} \sum (X_{1i} - \bar{X}_1)^2 & \sum (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) \\ \sum (X_{2i} - \bar{X}_2)^2 & \cdots \end{bmatrix}$$

نقبل H_0 اذا كانت

$$T^2 \leq T^2_{(critical)}$$

حرجة

where $T^2_{crit.} = \frac{p(n-1)}{n-p} F(\alpha, p, n-p)$

نرفض H_0 اذا كانت

$$T^2 > T^2_{crit.}$$

اما فترات الثقة حول بيتا μ ، هي تلك الفترات التي تحقق المتباينة التالية:

$$n(\bar{X}-\mu^*)' S^{-1} (\bar{X}-\mu^*) \leq T^2_{crit.}$$

Example

A r.s of size (101) taken from $N_2(\mu, \Sigma)$ and unbiased estimators for μ, Σ are

$$\underline{\bar{X}} = (5.524 \quad 34.97)'$$

$$S = \begin{bmatrix} 200.54 & 126.99 \\ & 119.68 \end{bmatrix}$$

① Test the hypothesis that the sample taken from $N_2 \sim \left[\begin{pmatrix} 60 \\ 50 \end{pmatrix}, \Sigma\right]$, $\alpha = 0.05$

$H_0 : \underline{\mu} = \begin{bmatrix} 60 \\ 50 \end{bmatrix}$ vs $H_1 : \underline{\mu} \neq \begin{bmatrix} 60 \\ 50 \end{bmatrix}$

$T^2 = n(\underline{\bar{X}} - \underline{\mu}_0)' S^{-1} (\underline{\bar{X}} - \underline{\mu}_0)$

Scanned with CamScanner

$$= (101)\begin{bmatrix} 55.24 - 60 \\ 74.97 - 50 \end{bmatrix}' \begin{bmatrix} 200.54 & 126.99 \\ & 119.68 \end{bmatrix} \begin{bmatrix} 55.24 - 60 \\ 34.97 - 50 \end{bmatrix}$$

$\therefore T^2 = 357.34$ $F(\alpha, P, n-P)$

$$T^2_{crit.} = \frac{P(n-1)}{n-P} F_{(0.05,\ 2,\ 99)} = \frac{2(101-1)}{101-2}$$

↑ ناخذ صيت ∞ نقوض بـ 99

$$= \frac{200}{99}(4.83) = 9.758 \quad F_{(0.05,\ 2,\ \infty)}$$

$\therefore T^2 > T^2_{crit.} \Rightarrow$ reject H_0 رفض

ملاحظة: اذا القيمة الحرجة T^2 حاذي (١) جاز نقبل (H_0) بدون حتى الكتابة

(2) Do you think that the vector $(50\ \ 50)'$ lie inside the confidence interval ?

$$n(\bar{x} - \mu^*)' S^{-1} (\bar{x} - \mu^*) \leq T^2_{crit.}$$

$$(101)\begin{bmatrix} 55.24 - 50 \\ 34.97 - 50 \end{bmatrix}' \begin{bmatrix} 200.54 & 126.99 \\ & 119.68 \end{bmatrix}^{-1} \begin{bmatrix} 55.24 - 50 \\ 34.97 - 50 \end{bmatrix}$$

$$\leq 9.75$$

Scanned with CamScanner