教育학석사학위請求論文

t-分佈에 관한 연구

A Note on the t-distribution

1987年 8月

指導教授 具 振興

이論文을 석사학위論文으로 提出함

仁荷大學校 教育大學院

數學教育 專攻

金 銀 淑
本論文은 金銀淑의 碩士學位論文으로 認准함

1987年 8月

主審 金炳文

副審 吳滋興

副審 曹仲鉉
謝 辭

내 안의 빈 공간에 무언가 가득히 담아보고 싶었습니다. 結婚이라는 인생에 젖어보고 새 생명 나의 딸 jemma도 얻었습니다. 보잘것 없지만 이렇게 얕은 지식도 모아봅니다. 갈증은 계속 갈 증을 이루고 빈 공간은 아직도 하진해 보이지만 언제가 넘쳐흘러 이웃에게 도움되려 맡습니다.

本論文가 완성되기까지 指導해 주신 具滋興 教授님을 비롯하여 數學科 教授님들께 깊은 感謝를 드립니다.

끝으로 공부할 수 있도록 많은 배려와 힘을 주신 北仁川女中의 여러 先生님들께도 感謝를 드립니다.

1987生 8月

金 銀 淑
<table>
<thead>
<tr>
<th>章节</th>
<th>标题</th>
<th>页码</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>序 論</td>
<td>1</td>
</tr>
<tr>
<td>II.</td>
<td>定義 及 範疇性</td>
<td>2</td>
</tr>
<tr>
<td>III.</td>
<td>本 論</td>
<td>7</td>
</tr>
<tr>
<td>III-1</td>
<td>t - 分布與 概率 基本定理</td>
<td>7</td>
</tr>
<tr>
<td>III-2</td>
<td>實 例</td>
<td>18</td>
</tr>
<tr>
<td>IV.</td>
<td>結 論</td>
<td>23</td>
</tr>
<tr>
<td>參考文獻</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>國文要約</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>英文抄錄</td>
<td></td>
<td>26</td>
</tr>
</tbody>
</table>

203520
I. 序論

正規母集団 \( N(\mu, \sigma^2) \) 에서 홍부한 평균 \( \bar{X} \) 의 분포를 살펴는 데 있어서 분산 \( \sigma^2 \) 이 미분한 경우 대표본에서는 평균 분산 \( \hat{\sigma}^2 \) 으로 대용할 수 있으나 표본 크기 \( n \) 이 작은 경우 (\( n \leq 30 \)) \( \hat{\sigma}^2 \) 의 산출 (표본 변동)가 크므로 분산 \( \sigma^2 \) 을 평균 분산 \( \hat{\sigma}^2 \) 로 대용할 수 없다. 이러한 문제를 해결하기 위하여 고셋트 (William S. Gosset, 1876〜1936) 는 분산 \( \sigma^2 \) 을 포함하지 않는 통계량

\[ T = \frac{\bar{X} - \mu}{\left( \frac{\hat{\sigma}}{\sqrt{n-1}} \right)} = \frac{\sqrt{n-1}}{\hat{\sigma}} \left( \bar{X} - \mu \right) \]

의 분포를 고안하여 "Student의 \( t \)-분포" 라 하였다. 이것이 표본 분포의 계기의 기회가 되었다.

본 연구에서는 \( t \)-분포의 확률 분포의 평균, 분산 및 성질을 연구하여 본으로써 표본 분포에 관한 이해를 도모하고 한다.

-1-
## Ⅱ. 定義 및 보조정리

(정의 1)

確率變數 \( X \) 의 確率密度函數 \( f(x) \) 가

\[
f(x) = \frac{1}{\sqrt{2\pi \sigma}} e^{-\left(x - \mu \right)^2 / 2\sigma^2} \quad (-\infty < x < \infty)
\]

일 때, 確率變數 \( X \) 는 平均이 \( \mu \)이고 分散이 \( \sigma^2 \)인 正規分佈 \( N(\mu, \sigma^2) \)을 따른다고 한다.

(정의 2)

確率變數 \( X \) 가 麥克反曲 \( \alpha, \beta (\alpha > 0, \beta > 0) \)를 가지는 確率密度函數

\[
f(x) = \begin{cases} 
\frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} & (0 < x < \infty) \\
0 & (x \leq 0)
\end{cases}
\]

를 가진 때 確率變數 \( X \) 는 麥克反曲 \( \alpha, \beta \)를 가지는 감마分佈 \( \Gamma(\alpha; \beta) \)를 따른다고 한다.

(정의 3)

(정의 2)에서 \( \alpha = \frac{\gamma}{2}, \beta = 2 \)일 때는 確率變數 \( X \) 는 카이제곱分布 \( (\chi^2\text{-distribution}) \)를 따른다. 특히 \( \alpha = 1 \)일 때 確率變數 \( X \)는
지수 분포 (exponential distribution) 를 따르다고 한다.

(정의 4)

확률 변수 $Z$ 는 표준 정규 분포 $N(0, 1)$ 을 따르고 이와는 [독립인] 확률 변수 $V$ 가 카이제곱 분포를 따를 때, 확률 변수 $T$ 가 $T = Z / \sqrt{V/k}$ 로 정의되면 $T$ 는 자유도 (degree of freedom) 가 $k$ 인 $t$ 분포를 따르다고 한다.

(정의 5)

확률 변수 $X_1, X_2$ 의 [결합] 밀도 함수를 $f(x_1, x_2)$ 라고 하면 $X_1, X_2$ 의 각각의 [周邊] 밀도 함수 (marginal probability density function) 는 다음과 같다.

$$f_1(x_1) = \begin{cases} \sum f(x_1, x_2) & (X_1 \textit{ 이 } \text{연속인 } \text{경우}) \\ \int_{-\infty}^{x_1} f(x_1, x_2) \, dx_2 & (X_1 \textit{ 이 } \text{연속인 } \text{경우}) \end{cases}$$

$$f_2(x_2) = \begin{cases} \sum f(x_1, x_2) & (X_2 \textit{ 가 } \text{연속인 } \text{경우}) \\ \int_{-\infty}^{\infty} f(x_1, x_2) \, dx_1 & (X_2 \textit{ 가 } \text{연속인 } \text{경우}) \end{cases}$$

(정의 6)

확률 변수 $X_1, X_2$ 의 결합 밀도 함수가 각각의 [周邊] 밀도 함수의 곱과 같으면, 즉 $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$ 가 성립하면 확률 변수 $X_1, X_2$ 는 확률적으로 [独立] (stochastically independent)이다.
(보조정리 1)

확률 변수 \(X\)가 정규 분포 \(N(\mu, \sigma^2)\)을 따를 때 \(Z = \frac{X - \mu}{\sigma}\)로 정의된 확률 변수 \(Z\)는 평균이 0이고 분산이 1인 표준 정규 분포 \(N(0, 1)\)을 따른다.


(보조정리 2)

확률 변수 \(X\)가 정규 분포 \(N(\mu, \sigma^2)\)을 따를 때 \(V = (X - \mu)^2/\sigma^2\)으로 정의된 확률 변수 \(V\)는 자유도가 1인 카이제곱 분포 \(\chi^2(1)\)을 따른다.


(보조정리 3)

확률 변수 \(X_1, X_2, \ldots, X_n\)이 각각 정규 분포 \(N(\mu, \sigma^2)\)을 따르며 서로 독립인 때 \(Y = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2}\)으로 정의된 확률 변수 \(Y\)는 자유도가 \(n\)인 카이제곱 분포 \(\chi^2(n)\)을 따른다.


(보조정리 4)

임의의 실수 \(\alpha > 0\)에 대한 감마 함수 (gamma function)은

\[
\Gamma(\alpha) = \int_{0}^{\infty} x^{\alpha-1} e^{-x} \, dx
\]

이다.
이 때 $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$, 즉 $\Gamma(\alpha) = (\alpha - 1)!$ 이다.

(증명) [1]의 103 p.

(보조정리 5)

任意의 실수 $\alpha > 0, \beta > 0$에 대하여

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1 - x)^{\beta-1} \, dx$$

로 정의되는 $\alpha, \beta$의 함수 $B(\alpha, \beta)$를 베타함수 (Beta function)이라 한다.

이 때 다음 관계식이 성립한다.

$$B(\alpha, \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

(증명) [1]의 104 p.

(보조정리 6)

中心極限定理 (Central limit theorem)

모집단의 확률변수 $X$가有限平均 $\mu$ 및 분산 $\sigma^2$를 가지는 경우 표본크기 $n$이 표본의 통계량 $\bar{X} (= \frac{1}{n} \sum_{i=1}^{n} x_i)$의 분포는 표본크기 $n$이 상당히 클 때 近似正規分布 (asymptotic normal distribution) $N(\mu, \frac{\sigma^2}{n})$을 따름다.

(보조정리 7) 카이제곱분포의 加法性

확률변수 \( X_1, X_2, \ldots, X_n \) 이 서로 독립이며 각각 자유도가 \( k_1, k_2, \ldots, k_n \)인 카이제곱분포 \( \chi^2(k_1), \chi^2(k_2), \ldots, \chi^2(k_n) \)을 따른다면 \( Y = X_1 + X_2 + \cdots + X_n \)로 정의된 확률변수 \( Y \)는 자유도가 \( k_1 + k_2 + \cdots + k_n \)인 카이제곱분포 \( \chi^2(k_1 + k_2 + \cdots + k_n) \)을 따른다.

Ⅲ. 本 論

中心極限定理 (보조정리 6) 으로부터 우리는 확률 변수 \( X_1, X_2, \ldots, X_n \)이 정규 모집단 (normal population) \( N(\mu, \sigma^2) \) 으로부터의 확률標
本일 때 표본 평균 \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i \) 는 평균이 \( \mu \)이고, 분산이 \( \frac{\sigma^2}{n} \)인 정규 분포 \( N(\mu, \frac{\sigma^2}{n}) \)을 따르며 표준화 변수 \( \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \)는 평균이 0이고, 분산이 1인 표준정규분포 (Standard normal distribution) \( N(0, 1) \)을 따름을 알았다. 그런데 평균 \( \mu \)에 대한 통계적 추정에서 분산 \( \sigma^2 \)이 밝혀지지 않은 경우에는 \( \sigma^2 \) 대신 표본분산 \( s^2 \)을 대체하여 스튜다르트화 (Studented) 된 확률변수의 분포를 필요로 하는 경우가 많다. (참고문헌 [2] 144 p.) (정의 4)에서의 \( t \)-분포 (t-distribution)은 이러한 확률변수들의 분포를 잘 나타내어 준다.

Ⅲ-1. \( t \)-분포에 관한 기본정리

(정리 1)

(정의 4)에서 정의된 확률 변수 \( T = \frac{Z}{\sqrt{V/k}} \)의 확률밀도함수 (probability density function) \( h(t) \)는 다음과 같다.

\[
h(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right) \sqrt{\pi k}} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}, \quad -\infty < t < \infty \quad (1,1)
\]
여기서 \( h(t) \)는 자유도가 \( k \)인 \( t \)-분포의 확률밀도함수이다.

(증명)

\( Z \)와 \( V \)는 (정의 4)로부터 서로 독립인 확률변수이므로 (보조정리 6)으로부터 그들의 결합밀도함수는 각각의 확률밀도함수의 곱과 같다. 따라서 \( Z \)와 \( V \)의 결합밀도함수를 \( f(z,v) \)라 하면

\[
f(z,v) = f(z) \cdot f(v)
\]

이다.

그런데 (정의 4), (정의 1), (정의 3)으로부터

\[
f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty \quad (1, 3)
\]
\[
f(v) = \frac{1}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}} v^{\frac{k-1}{2}} e^{-\frac{v}{2}}, \quad 0 < v < \infty \quad (1, 4)
\]

그러므로

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

\[(-\infty < z < \infty, \ 0 < v < \infty) \quad (1, 5)\]

이다.

다음에 확률변수 \( W = V \)를 정의하자.
즉,

\[
T^* = \frac{Z}{\sqrt{V/k}} \tag{1,6}
\]

\[W = V\]

이러 정의하면 (1,6) 으로부터

\[
\begin{cases}
z = \frac{t\sqrt{v}}{\sqrt{k}} \\
v = w
\end{cases} \tag{1,7}
\]

여기서 \(t, z, v, w\) 는 \(T, Z, V, W\) 의 일차원이며, 이 변환은 \(\{(z, v) \mid -\infty < z < \infty, 0 < v < \infty\}\) 를 \(\{(t, w) \mid -\infty < t < \infty, 0 < w < \infty\}\) 에로의 1:1 변환 (one-to-one transformation)이며, 야코비안 (Jacobian) \(J\) 는

\[
J = \begin{vmatrix}
\frac{\partial z}{\partial t} & \frac{\partial z}{\partial w} \\
\frac{\partial v}{\partial t} & \frac{\partial v}{\partial w}
\end{vmatrix} = \begin{vmatrix}
\frac{\sqrt{w}}{\sqrt{k}} & \frac{t}{2\sqrt{kW}} \\
0 & 1
\end{vmatrix} = \frac{\sqrt{w}}{k} \tag{1,8}
\]

이므로 확률 변수 \(T\) 와 \(W\) 의 결합 밀도 함수를 \(g(t, w)\) 라 하면

\[
g(t, w) = f\left(\frac{t\sqrt{w}}{\sqrt{k}}, w\right) |J|
\]
\[ f(t) \sqrt{\frac{w}{k}} \cdot f(w) | J | \]

\[ = \frac{1}{\sqrt{2\pi}} e^{-\frac{(t \sqrt{\frac{w}{k}})^2}{2}} \cdot \frac{1}{\Gamma\left(\frac{k}{2}\right)^{2^{\frac{k}{2}}}} w^{k-1} \cdot e^{-\frac{w}{2}(1+\frac{t^2}{k})} \cdot \sqrt{\frac{w}{k}} \]

\[ = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma\left(\frac{k}{2}\right)^{2^{\frac{k}{2}}}} w^{k-1} \cdot e^{-\frac{w}{2}(1+\frac{t^2}{k})} \cdot \left(\frac{w}{k}\right)^{\frac{1}{2}} \]

(\(-\infty < t < \infty, \ 0 < w < \infty\))

(정의 5)에 의하여 \( w \)에 관하여 적분하여 \( t \)에 관한 \( \text{確率密度} \)

\( h(t) \)를 구하면,

\[ h(t) = \int_0^\infty g(t, w) \, dw \]

\[ = \int_0^\infty \frac{1}{\sqrt{2\pi k}} e^{-\frac{w^{k+1-1}}{2} \cdot e^{-\frac{w}{2}(1+\frac{t^2}{k})}} \, dw \]

\[ = \frac{1}{\sqrt{2\pi k}} \int_0^w w^{k+1-1} \cdot e^{-\frac{w}{2}(1+\frac{t^2}{k})} \, dw \quad (1,10) \]

\[(1,10) \quad \int_0^w w^{k+1-1} \cdot e^{-\frac{w}{2}(1+\frac{t^2}{k})} \, dw \quad \text{의 값을 구하기 위하여} \]

\[ A = -\frac{w}{2} (1 + \frac{t^2}{k}) \quad \text{로 치환하면} \]

-10-
\[ w = \frac{2A}{(1 + \frac{t^2}{k})} \cdot \text{고} \quad dw = \frac{2}{(1 + \frac{t^2}{k})} \cdot dA \]

따라서

\[
\int_0^\infty w^{\frac{k+1}{2}-1} \cdot e^{-w(1 + \frac{t^2}{k})} \cdot \frac{2A}{(1 + \frac{t^2}{k})} \cdot e^{-A} \cdot \frac{2}{(1 + \frac{t^2}{k})} \cdot dA
\]

\[
= \int_0^\infty \left( \frac{2A}{(1 + \frac{t^2}{k})} \right)^{\frac{k+1}{2}} \cdot A^{\frac{k+1}{2}-1} \cdot e^{-A} \cdot dA
\]

\[
= \left( \frac{2}{(1 + \frac{t^2}{k})} \right)^{\frac{k+1}{2}} \int_0^\infty A^{\frac{k+1}{2}-1} \cdot e^{-A} \cdot dA
\]

\[
= \Gamma\left( \frac{k+1}{2} \right) \quad (1,11)
\]

그렇게 (보조정리 4)에 의하여

\[
\int_0^\infty A^{\frac{k+1}{2}-1} \cdot e^{-A} \cdot dA = \Gamma\left( \frac{k+1}{2} \right) \quad (1,12)
\]

이므로

\[
(1,11) = \left( \frac{2}{(1 + \frac{t^2}{k})} \right)^{\frac{k+1}{2}} \cdot \Gamma\left( \frac{k+1}{2} \right) \quad (1,13)
\]

\[
\text{이(1,13)을 (1,10) 식에 대입하면}
\]

-11-
\[ h(t) = \frac{1}{\sqrt{2\pi k}} \frac{1}{\Gamma\left(\frac{k+1}{2}\right)} \left(\frac{2}{1 + \frac{t^2}{k}}\right)^{\frac{k+1}{2}} \cdot \Gamma\left(\frac{k+1}{2}\right) \]

\[ = \frac{\frac{k+1}{2}}{\sqrt{2\pi k} \Gamma\left(\frac{k}{2}\right) 2^k} \cdot \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}} \]

\[ = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}, \quad (-\infty < t < \infty) \quad (1.14) \]

을 얻는다.

(정리 2)

自由度가 \( k \)인 \( t \) 분포의 평균 \( E(t) \), 분산 \( V(t) \)는 다음과 같다.

i) \[ E(t) = 0 \quad (1.15) \]

ii) \[ V(t) = \frac{k}{k-2} \quad (1.16) \]

(증명)

(i) (정리 1)로부터

\[ h(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}, \quad (-\infty < t < \infty) \]
이다.

그런데 $h(t)$는 우함수 (even function)이므로

$$E(t) = \int_{-\infty}^{\infty} t \cdot h(t) \, dt = 0$$  \hspace{1cm} (1,17)

(ii) $E(t) = 0$ 이므로

$$V(t) = E(t^2) - [E(t)]^2 = E(t^2)$$

$$= \int_{-\infty}^{\infty} t^2 h(t) \, dt$$

$$= \int_{-\infty}^{\infty} t^2 \frac{\Gamma \left( \frac{k+1}{2} \right)}{\sqrt{\pi} k \Gamma \left( \frac{k}{2} \right)} \left( 1 + \frac{t^2}{k} \right)^{-\frac{k+1}{2}} \, dt$$

$$= \frac{\Gamma \left( \frac{k+1}{2} \right)}{\sqrt{\pi} k \Gamma \left( \frac{k}{2} \right)} \int_{-\infty}^{\infty} t^2 \left( 1 + \frac{t^2}{k} \right)^{-\frac{k+1}{2}} \, dt$$  \hspace{1cm} (1,18)

여기서 먼저 $\int_{-\infty}^{\infty} t^2 \left( 1 + \frac{t^2}{k} \right)^{-\frac{k+1}{2}} \, dt$의 값을 구하기 위하여

$$(1 + \frac{t^2}{k})^{-1} = v$$라고 치환하면 $t^2 = \frac{k(1 - v)}{v}$이므로

$$dt = \frac{1}{2} \left[ \frac{k(1 - v)}{v} \right]^{-\frac{1}{2}} (-\frac{k}{v^2}) \, dv$$

그러므로

-13-
\[ \int_{-\infty}^{\infty} t^2 \left( 1 + \frac{t^2}{k} \right)^{-\frac{k+1}{2}} \, dt \]

\[ = 2 \int_{0}^{\infty} t^2 \left( 1 + \frac{t^2}{k} \right)^{-\frac{k+1}{2}} \, dt \]

\[ = 2 \int_{0}^{1} \frac{k(1-v)}{v} \cdot v^{\frac{k+1}{2}} \cdot \frac{1}{2} \left[ \frac{k(1-v)}{v} \right]^{-\frac{1}{2}} \left( -\frac{k}{v^2} \right) \, dv \]

\[ = \int_{0}^{1} \left[ \frac{k(1-v)}{v} \right]^{\frac{1}{2}} v^{\frac{k+1}{2}} \left( \frac{k}{v^2} \right) \, dv \]

\[ = k^{\frac{3}{2}} \int_{0}^{1} v^{\frac{k-4}{2}} (1-v)^{\frac{1}{2}} \, dv \]

\[ = k^{\frac{3}{2}} \int_{0}^{1} v^{\frac{k-2}{2} - 1} (1-v)^{\frac{3}{2} - 1} \, dv \quad (1.19) \]

그런데 (보조정리 5)로부터

\[ \int_{0}^{1} v^{\frac{k-2}{2} - 1} (1-v)^{\frac{3}{2} - 1} \, dv \]

\[ = B \left( \frac{k-2}{2}, \frac{3}{2} \right) \]

\[ = \frac{\Gamma \left( \frac{k-2}{2} \right) \Gamma \left( \frac{3}{2} \right)}{\Gamma \left( \frac{k+1}{2} \right)} \quad (1.20) \]

이므로

\[ \int t^2 \left( 1 + \frac{t^2}{k} \right)^{-\frac{k+1}{2}} \, dt \]
\[ V(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi} k \Gamma(\frac{k}{2})} \cdot k^2 \frac{\Gamma\left(\frac{k-2}{2}\right) \Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{k+1}{2}\right)} \]

\[ = k \frac{\Gamma\left(\frac{k-2}{2}\right) \Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{k-2}{2}\right) \Gamma\left(\frac{k-2}{2}\right)} \]

\[ = k \frac{\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi} \cdot (k-2)} \quad (\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}) \]

\[ = \frac{k}{k-2} \quad (k \geq 3) \quad (1, 22) \]

을 얻는다.

(정리 3)

確率変数 \( X_1, \ldots, X_n \) 이 서로 독립이며 각각 평균이 \( \mu \)이고 분산이 \( \sigma^2 \)인 정규 분포 \( N(\mu, \sigma^2) \)을 따른다면, 통계량 \( \frac{nS^2}{\sigma^2} \)은 자유도가 \( n-1 \)인 카이제곱 분포 \( \chi^2(n-1) \)을 따른다.
여기서

\[ S^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2 \]

(증명)

\[ nS^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=1}^{n} (X_i - \mu + \mu - \overline{X})^2 \]

\[ = \sum_{i=1}^{n} \{ (X_i - \mu)^2 + 2 (X_i - \mu) (\mu - \overline{X}) + (\mu - \overline{X})^2 \} \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 + 2 (\mu - \overline{X}) \sum_{i=1}^{n} (X_i - \mu) + n (\mu - \overline{X})^2 \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 + 2 (\mu - \overline{X}) (\sum_{i=1}^{n} X_i - n \mu) + n (\mu - \overline{X})^2 \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 + 2 (\mu - \overline{X}) (n\overline{X} - n \mu) + n (\mu - \overline{X})^2 \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 - 2n (\mu - \overline{X})^2 + n (\mu - \overline{X})^2 \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 - n (\overline{X} - \mu)^2 \quad (1, 23) \]

그러므로

\[ \frac{nS^2}{\sigma^2} = \frac{\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} - n \frac{(\overline{X} - \mu)^2}{\sigma^2}}{\sigma^2} \]

\[ = \frac{\sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2 - \left( \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \right)^2}{\sigma^2} \quad (1, 24) \]

그러면 보조정리 2와 3으로부터

-16-
\[ \sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2 \text{은 自由度가 } n \text{인 카이제곱분포 } \chi^2(n) \text{을 따르고} \]

\[ \left( \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \right)^2 \text{은 自由度가 } 1 \text{인 카이제곱분포 } \chi^2(1) \text{을 따른다.} \]

그러므로 (보조정리 7 )에 의하여 \( \frac{nS^2}{\sigma^2} \)은 自由度가 n-1인 카
이제곱분포 \( \chi^2(n-1) \)을 따른다.

(정리 4)

確率變數 \( X_1, X_2, \ldots, X_n \)이 서로 獨立이며 각각 平均이 \( \mu \)이고
分散이 \( \sigma^2 \)이 정규分布 \( N(\mu, \sigma^2) \)을 따를 때 \( t \)-검정통계량 (\( t \)-test statistic)

\[ T = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}} \]

는 自由度 \( k = n - 1 \)인 \( t \)-분포를 따른다.

여기서

\[ \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \]

(증명)

\[ T = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}} \]
\[
\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{\hat{\sigma}}{\sqrt{n}} \frac{\hat{\sigma}}{\sigma} \\
= \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sqrt{\frac{(n-1) \hat{\sigma}^2}{\sigma^2}} \frac{1}{n-1} \\
= \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sqrt{\frac{nS^2}{\sigma^2}} \frac{1}{(n-1)}
\]

그러면 중심극한정리(보조정리 6)와 (정리 3)로부터 \( \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \) 은

평균이 0이고 분산이 1인 표준정규분포 \( \mathcal{N}(0, 1) \)을 따르고

\( \frac{nS^2}{\sigma^2} \)은 자유도가 \( n-1 \)인 카이제곱분포 \( \chi^2(n-1) \)을 따르므로

(정의 4)에 의해 \( T = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}} \)은 자유도가 \( n-1 \)인 \( t \)-분포를 따른다.

\( \text{III-2. 실례} \)

(i) 확률변수 \( X_1, X_2, \cdots, X_n, X_{n+1} \)이 서로 독립이며 정규분포

\( \mathcal{N}(\mu, \sigma^2) \)을 따를 때,

\[
\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{이상}
\]

이러 나머지 통계량

\( \text{고려 18} \)
\[ L = \frac{C (\overline{X} - X_{n+1})}{S} \]  \hspace{1cm} (2,2)

가 \( t \)-분포를 따를 상수 \( C \) 를 구하고, 아울러 \( n \)이 8일 경우

\[ P_r (\overline{X} - kS < X_r < \overline{X} + kS) = 0.80 \]  \hspace{1cm} (2,3)

을 만족하는 \( k \) 를 구하여 보자.

(정리 3)에 의하여 \( \frac{nS^2}{\sigma^2} \) 은 자유도 \( k = n - 1 \)인 카이제곱분포 \( \chi^2 (n - 1) \) 을 따르며 (보조정리 6)에 의하여 \( \overline{X} \)는 정규 분포 \( N (\mu, \frac{\sigma^2}{n}) \)을 따른다.

아울러 \( \overline{X} - X_{n+1} \)은 정규 분포 \( N (0, \sigma^2 + \frac{\sigma^2}{n}) \)을 따른다.

그리므로 \( t \)-분포의 정의에 따라

\[
\frac{\overline{X} - X_{n+1}}{\sqrt{\sigma^2 + \frac{\sigma^2}{n}}} \left/ \sqrt{\frac{nS^2}{\sigma^2}} \right/ \frac{n-1}{n-1}
\]  \hspace{1cm} (2,4)

는 자유도 \( k = n - 1 \)인 \( t \)-분포 \( t(n-1) \)을 따른다.

\[ (2,4) = \frac{\overline{X} - X_{n+1}}{\sqrt{\frac{\sigma^2 (n+1)}{n}}} \left/ \sqrt{\frac{nS^2}{\sigma^2 (n-1)}} \right/ \frac{nS^2}{\sigma^2 (n-1)} \]

\[ = \frac{\overline{X} - X_{n+1}}{\sqrt{\frac{\sigma^2 (n+1)}{n}}} \sqrt{\frac{nS^2}{\sigma^2 (n-1)}} \]

\[ = \frac{\overline{X} - X_{n+1}}{\sqrt{\frac{n+1}{n-1} \cdot S^2}} \]

-19-
\[ = \sqrt{\frac{n-1}{n+1}} \cdot \frac{\overline{X} - X_{n+1}}{S} \]  
(2.5)

그러므로

\[ C = \sqrt{\frac{n-1}{n+1}} \]  
(2.6)

이다.

\[ n = 8 \text{이면 } t_{0.9}(7) = 1.415 \text{ ( } t \text{-분포표에 의하여) 이므로} \]

\[ P_r \left( -1.415 < \sqrt{\frac{7}{9}} \cdot \frac{\overline{X} - X_9}{S} < 1.415 \right) = 0.80 \]  
(2.7)

\[ P_r \left( \overline{X} - 1.415 \sqrt{\frac{7}{9}} \cdot S < X_9 < \overline{X} + 1.415 \sqrt{\frac{7}{9}} \cdot S \right) \]

\[ = 0.80 \]  
(2.8)

따라서

\[ k = 1.415 \times \sqrt{\frac{7}{9}} = 1.604 \]  
(2.9)

이다.

(ii) 학생 9명이 1, 2학기 수학성적이 다음 표와 같은 때

\[ \mu_1 = \mu_2 \text{의 95\% } \text{信頼區間을 구하여 보자.} \]
<table>
<thead>
<tr>
<th>성적</th>
<th>번호</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>학 기</td>
</tr>
<tr>
<td>2</td>
<td>학 기</td>
</tr>
<tr>
<td>$W_i = X_{1i} - X_{2i}$</td>
<td>6</td>
</tr>
</tbody>
</table>

$$W_i = W_2 = W_4 = 6$$
$$W_3 = -2$$
$$W_5 = 7$$
$$W_6 = 3$$
$$W_7 = W_8 = 1$$
$$W_9 = 5$$

이므로

$$\bar{W} = \frac{1}{n} \sum_{i=1}^{10} W_i = \frac{1}{9} \times 33 = 3.6$$  \hspace{1cm} (2.11)

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (W_i - \bar{W})^2 = \frac{1}{n-1} \sum_{i=1}^{n} W_i^2 - \frac{n}{n-1} \bar{W}^2$$

$$= \frac{1}{8} \times 197 - \frac{9}{8} \times (3.6)^2 \approx 9.5$$ \hspace{1cm} (2.12)

따라서 信頼限界是

$$\bar{W} \pm t_{0.05} (8) \frac{\hat{\sigma}}{\sqrt{n}} \approx 3.6 \pm 0.306 \times \frac{3.08}{3}$$
\[ \pm 3.6 \pm 2.369 \]  

(2.13)

그러므로 \( E(\overline{W}) = \mu_1 - \mu_2 \)의 信頼係數 95\% 인 信頼區間은

\[ 1.297 \leq \mu_1 - \mu_2 \leq 6.035 \]  

(2.14)
IV. 结 論

$t$-분포에 대한 임의의 구간, 개별 범위를 구하려면 분산 $\sigma^2$이 미지인 경우라도 항상 이용할 수 있는 것은 아니다.

$t$-분포에 의한 임의의 구간의 관리는 확률변수 $X$가 정규분포를 따르는 가정을 필요로 하고 있다.

따라서 $X$가 적어도 정규분포에 따르는 보장이 없는 한 $t$-분포를 적용할 수 없다.

본론에서의 예로는 $t$-분포의 이용범위를 다음과 같이 말할 수 있다.

(1) 원점단이 정규분포를 이룰 때, 이 분포로부터 임의 추출한 표본크기 30 미만의 임의 평균의 분포에서 $t$-분포가 이용된다.

(2) 모분산이 미지인 경우, $\bar{X}$의 분포를 살펴 볼 때 $t$-분포를 이용하여 이 때 통계량

$$T = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}}, \quad \text{(단, } \hat{\sigma}^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2)$$

은 자유도 $k = n - 1$인 $t$-분포를 이룬다. 아울러 $t$-분포는 자유도 $k \geq 25$이면 정규분포에 근사하고 $k \to \infty$에서 표준정규분포

$N(0,1)$에 일치한다.
1. 具滋興，‘高級統計學’，仁荷大學校 出版部，1977.


要 約

任意의 標本을 基礎로 한 統計量을 하나의 確率變數로 보고 그에 持할 理論分布를 求하는 研究를 標本分布 (Sample distribution)에 持한 研究에 하며 이는 近代統計學의 中枢이론을 優성하고 있다.

特に, 小標本에 있어서 推定과 檢定, 分散分析에 자주 이용되는 分布로는 $X^2$ 分布, $t$ 分布, $F$ 分布 등이 있는데 本稿에서는 이들 分布 중 $t$ 分布에 持한 特性을 大勢하였다.
ABSTRACT

The rapid development of statistics in recent years has brought about their increased application to problems in economic research, business, industry, agriculture, and many other fields.

Mathematical statistics depends heavily on the theory of sample distributions.

This paper is concerned with the properties and applications of Student's $t$-distribution.