碩士學位請求論文

Gamma 函數와 그 應用에 관한 研究

指導敎授 金 益 贊

濟州大學校 敎育大學院

數學敎育專攻

夫 淑 姬

1992年 8月

Gamma 函數와 그 應用에 관한 研究

指導教授 金 益 贊

이 論文을 教育學 碩士學位 論文으로 提出함

1992年 6月 日

濟州大學校 教育大學院 數學教育專攻

提出者 夫 淑 姫

夫淑姫의 教育學 碩士學位 論文을 認准함

1992年 7月 日

審査委員長 宋 錫 準

審 査 委 員 金 益 贊

審 査 委 員 金 鐵 洙

〈초 록〉

Gamma 函數와 그 應用에 관한 硏究

夫 淑 姫

濟州大學校 敎育大學院 數學敎育專攻

指導敎授 金 益 贊

본 논문에서는 n의 계승의 일반화로서 Gamma 함수의 발견에 대한 역사적 고찰에서 출발하여 이상적분으로서의 Gamma 함수의 수렴성과 주요 성질을 연구하였다. Gamma 함수에서 Gamma 분포 χ^2 분포를 유도하고 통계분야, 특히 추정 검정론에서 응용되고 있는 동일성과 독립성, 적합도 검정등에 실제 활용되고 있는 예를 보임으로서 Gamma 함수에 대한 중요성을 재확인하고자 하였다.

目 次

I. 序　　論

n이 양의정수가 아닌 경우에도 계승 $n!$을 일반화하여 보려는 많은 시도가 Gamma 함수 발견의 발단이라고 알려지고 있다. 본 논문에서는 Gamma함수의 역사적 발견의 의의를 고찰하며 이러한 발견과 정의과정을 음미하고 나아가서는 Gamma함수의 성질과 특성을 조사한다. 또 이 함수에 대한 이해를 넓히고 이 함수가 응용되는 분야, 특히 통계분야에서의 활용, 이를테면 지수분포와 χ^2-분포, Beta분포에 의한 추정과 가설검정 및 적합도 검정분야 등에 대하여 그 이론과 실제 응용관계를 고찰함으로써, 고등미적분 및 통계분야에서의 Gamma함수에 대한 중요성을 인식시키고, 또 이를 요약, 정리하여 보려는 것이 본 논문의 목적이다.

n 이 양의정수가 아닐때 $n!$을 일반화하려는 최초의 시도는 L.Euler(1707 ~ 1783)가 C.Goldbach(1690 ~ 1764) 에게 보낸 편지라고 알려지고 있다. 즉 *Euler*는 다음과 같은 계수를 변형하여 無限乘積形의 식을 생각하였다.

n, m을 양의정수라고 할 때

$$(n-1)! = \frac{1\cdot 2\cdots(m-1)\cdot m^n}{n(n+1)(n+2)\cdots(n+m-1)} \times \left[\frac{m}{m}\cdot\frac{m+1}{m}\cdot\frac{m+2}{m}\cdots\frac{m+n-1}{m}\right]$$

여기서 n을 고정하고 $m \to \infty$라고 하면 위 식의 오른쪽 [ ]은 1에 수렴함으로

(1-1) $$(n-1)! = \lim_{m \to \infty} \frac{(m-1)! \cdot m^n}{n(n+1)(n+2)\cdots(n+m-1)}$$

이 된다. 한편 (1-1)의 우변식은 n이 정수가 아니라도 의미를 가지게 되므로 이 극한치에 의하여 계승의 일반화가 가능하며 이 극한은 解析的 階乘이라고 불렀다. 다시 (1-1)의 우변에서

$$\begin{aligned} m^n &= \left(\frac{2}{1} \cdot \frac{3}{2} \cdots \frac{m}{m-1}\right)^n \\ &= \left(1+\frac{1}{1}\right)^n \left(1+\frac{1}{2}\right)^n \cdots \left(1+\frac{1}{m-1}\right)^n \end{aligned}$$

이 되므로

(1-2) $$(n-1)! = \frac{1}{n} \prod_{m=1}^{\infty} \left(1+\frac{1}{m}\right)^n \left(1+\frac{n}{m}\right)^{-1}$$

형으로 *Euler*는 표시하였다. 현대에 와서 Gamma함수는 *Laplace* 변환에 의한 적분

(1-3) $$\Gamma(k) = \int_0^\infty e^{-t} t^{k-1}\, dt$$

에 의해서 정의된다. k가 양의정수 n일 때 (1-3)식은 $(n-1)!$이 된다.

한편 (1-3)에서 $e^{-t} = u$로 치환하면

(1-4) $$\Gamma(k) = \int_0^1 \left(ln\frac{1}{u}\right)^{k-1} du$$

로 표시된다.

역사적으로 볼 때 적분에 의한 Gamma함수의 정의는 먼저 (1-4)의 형태로 $Euler$ 에 의하여 도입되었다. 즉 $Euler$ 는 (1-3), (1-4)식에서 주로 유리수 k 에 대한 값으로 사용하였다. (1-3)식이 k 의 함수라고 생각하고 Gamma함수라 명명하면서 $\Gamma(k)$의 기호로 나타낸 것은 $A.M.Legendre$(1752 ~ 1833)로 알려지고 있다. 특히 $Legendre$ 는 $Beta$ 함수

(1-5) $$\beta(m,n) = \int_0^1 t^{m-1}(1-t)^{n-1}\,dt$$

를 나타내는 적분을 제1종의 $Euler$ 적분, (1-3)식을 제2종의 $Euler$ 적분이라고 불렀다.

한편 $C.F.Gauss$(1777 ~ 1855) 는 $n!$ 의 일반화로서 $k! = \Gamma(k+1)$ 또는 $k! = \prod(k)$라는 기호를 사용하였던 바 현재는 모두 $\Gamma(k)$ 의 기호를 사용하고 있다. 한편 (1-5)가

(1-6) $$\beta(m,n) = \frac{\Gamma(m)\cdot\Gamma(n)}{\Gamma(m+n)}$$

으로 표시된다는 사실은 Euler에 의해 발견되었던 바, 여기서 (1-1) 또는 (1-2)식이 (1-3)과 같은 값을 갖는다는 것에 대해서는 (1-3)에서 점화식

(1-7) $$\Gamma(k+1) = k\Gamma(k)$$

가 유도되고 이것을 반복하고 잉여항을 생각하면 $0 \le k \le 1$ 일 때

$$\Gamma(k) = \frac{1\cdot 2\cdots(m-1)}{k(k+1)\cdots(k+m-1)}m^k(1+\theta\cdot\varepsilon_m)$$

여기서 $-1 < \theta < 1$, $\varepsilon_m \to 0$에 의해 $\Gamma(k)$는 $0 < k \leq 1$에서 (1-1)과 같음을 확인할 수 있다. 한편 (1-1) 및 (1-3)도 점화식 (1-7)을 만족시키게 됨을 확인하면 된다. 따라서 $Euler$의 常數

$$\gamma = \lim_{m\to\infty}\left(1+\frac{1}{2}+\cdots\frac{1}{m}\right)-\log m$$

은 역사적으로 $\Gamma(k)$의 無限乘積 표시와 관련되어 나타난 것임을 알 수 있다.

Ⅱ. Gamma 함수와 그 성질

1. Gamma 함수

무한급수와 異常積分 사이의 유사성에 비추어 멱급수

(2-1) $$G(x) = \sum_{n=0}^{\infty} f(n)x^n$$

을

(2-2) $$\int_0^{\infty} f(t)x^t \, dt$$

로 생각하고 x를 e^{-s}으로 치환함으로써 $f(t)$의 $Laplace$의 변환인

(2-3) $$F(s) = \int_0^{\infty} f(t)e^{-st} \, dt$$

를 정의한다. $G(x)$의 수렴반경을 생각하듯이 (2-2)가 어떤 γ^*에 대하여 $0 \le x < \gamma^*$에서 수렴한다면 $F(s)$도 $0 < e^{-s} < \gamma^*$에서 수렴한다. 이제 (2-3)의 적분을 함수 $f(t)$가 $F(s)$로 변환하는 operator $\mathcal{L}$ 로 생각하여

(2-4) $$\mathcal{L}(f) = F$$

로 표시하자.

$f(t) = t^k$으로 두었을 때의 $Laplace$의 변환 :

$$\int_0^\infty t^k e^{-st}\,dt \tag{2-5}$$

를 생각하자. 단, $t = 0$에서 이 적분이 발산함을 피하기 위하여 매개변수 k는 -1보다 크다고 하자. 만일 $k = 0$이거나 또는 양의정수이면

$$\int_0^\infty e^{-st}\,dt = \frac{1}{s}, \qquad \int_0^\infty te^{-st}\,dt = \frac{1}{s^2}, \cdots, \cdots, \qquad (s > 0) \tag{2-6}$$

이 되고 일반적인 경우가 부분적분법에 의해

$$\int_0^\infty t^k e^{-st}\,dt = \frac{k}{s}\int_0^\infty t^{k-1}e^{-st}\,dt \quad : \quad s > 0 \tag{2-7}$$

가 됨을 확인할 수 있다. 이제 귀납법을 적용함으로써

$$\mathcal{L}[t^k] = \int_0^\infty t^k e^{-t}\,dt = \frac{k}{s}\cdot\frac{k-1}{s}\cdots\frac{1}{s}\cdot\frac{1}{s} = \frac{k!}{s^{k+1}}, \quad s > 0 \tag{2-8}$$

따라서 $s = 1$일 때

$$k! = \int_0^\infty e^{-t}t^k\,dt \tag{2-9}$$

가 되어 결국 계승(factorial)에 대한 일반화를 가능하게 한다. 즉, $k > -1$인 임의의 실수일 때 $k!$은 (2-9)에 의해 정의되고 k가 0 또는 양의정수이면

$$k! = \Gamma(k+1) = \begin{cases} 1, 2, \cdots k & : k > 0 \\ 1 & : k = 0. \end{cases} \tag{2-10}$$

로 정의됨으로써 Gamma함수가 유도된다.

정 의 1 :

$$\Gamma(k) = \int_0^\infty e^{-t} t^{k-1}\, dt \quad ; \quad k > 0 \tag{2-11}$$

2. Gamma 함수의 수렴성

Gamma 함수로 정의된 무한적분 (2-11)의 수렴성을 조사하기 위하여 다음의 몇 가지 예비정리를 생각하자.

예비정리 1. 만일 $x \geq a$일 때 $f(x)$가 연속이고 상수 A가 존재해서 $|f(x) \cdot x^k| < A$이고 $k > 1$이면 $\int_a^\infty f(x)\, dx$는 절대수렴한다. 또 만일 $|f(x) \cdot x^k| \geq A$이고 $k \leq 1$이면 발산한다.

증 명 :

i) $|f(x)| < \dfrac{A}{x^k}$ 이므로 $\displaystyle\int_a^\infty |f(x)|\, dx < A \int_a^\infty \frac{1}{x^k}\, dx,$

그런데 $k \neq 1$이면

$$\int_a^b \frac{1}{x^k}\, dx = \int_a^b x^{-k}\, dx = \frac{1}{1-k}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right)$$

또 $k > 1$ 이면

$$\lim_{b\to\infty} \frac{1}{1-k}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right) = \frac{1}{(k-1)a^{k-1}}$$

즉, $\int_a^x |f(x)|\,dx$는 x에 관한 증가함수이고 $\int_a^\infty f(x)\,dx$가 존재하므로

$A\int_a^\infty \frac{1}{x^k}\,dx$의 값이 존재한다.

$ii)$ $x \geq a$일 때 $|f(x)| \geq \frac{A}{x^k}$이니 따라서 $\int_a^b |f(x)|\,dx \geq A\int_a^b \frac{1}{x^k}\,dx$

만일 $|f(x)| \geq \frac{A}{x^k}$ 이면 $f(x)$가 연속이므로 $x \geq a$일 때 $f(x)$의 부호를 변경시킬 수 없다. 이제 $f(x) > 0$ 이라고 가정하면

$$\int_a^b |f(x)|\,dx = \int_a^b f(x)\,dx \geq A\int_a^b \frac{1}{x^k}\,dx$$

그런데

$$\int_a^b \frac{1}{x^k}dx = \begin{cases} \frac{1}{1-k}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right) & : k < 1 \\ logb - loga & : k = 1. \end{cases}$$

그리고 이것은 $b \to \infty$ 일 때 발산하므로 $\int_a^\infty f(x)\,dx$는 발산한다.

$f(x) < 0$의 경우(단, $x \geq a$)에는

$$\int_a^b |f(x)|\,dx = -\int_a^b f(x)\,dx$$

이고 $f(x) > 0$의 경우를 적용하면 발산한다.

예비정리 2. 만일 $k > 1$일 때 $\lim_{x\to\infty} x^k f(x) = L$ 이면 $\int_a^\infty f(x)dx$는 절대수렴한다. 또 $k \le 1$이고 $\lim_{x\to\infty} x^k f(x) = L \ne 0$이면 $\int_a^\infty f(x)dx$는 발산한다. 이 정리는 예비정리 1의 결과에 따라 쉽게 증명할 수 있다.

예 제 1. $\int_0^\infty \frac{1}{(1+x^6)^{1/3}} dx$ 일 때

$$\lim_{x\to\infty} x^2 \frac{1}{\left(\frac{1}{x^6}+1\right)^{1/3}} = \lim_{x\to\infty} \frac{1}{\left(\frac{1}{x^6}+1\right)^{1/3}} = 1$$

따라서 $\int_0^1 f(x)\,dx$ 는 수렴한다.

예비정리 3. 만일 함수 $f(x)$가 $0 < x \le b$에서 연속이고, $0 < k < 1$일 때 $|f(x)\cdot(x-a)^k| < A$가 되는 양수 A가 존재한다면 $\int_a^b f(x)dx$는 절대수렴한다. 만일 $k \ge 1$이고 $a < x \le b$일 때 $|f(x)(x-a)^k| > A$이면 $\int_a^b f(x)\,dx$는 발산한다.

예비정리 4. 만일 $0 < k < 1$이고 $\lim_{x\to a+}(x-a)^k f(x) = L$이면 $\int_a^b f(x)dx$는 절대수렴하고, $k \ge 1$이고 $\lim_{x\to a+}(x-a)^k f(x) = L \ne 0$이면 $\int_a^b f(x)dx$는 발산한다.

이제 Gamma 함수로 정의된 무한적분 (2-11)의 수렴성을 조사하여 보자. k가 $0 < k < 1$일 때는 (2-11)식의 피적분함수는 하한 $t = 0$에서 무한이 된다. 따라서 상한과 하한의 양쪽에서 적분값의 존재성을 조사하여야 한다. 즉 (2-11)식을 $t \geq 1$일 때는 연속이 되는 두 함수의 합 즉

(2-12) $$\Gamma(k) = \int_0^1 e^{-t}t^{k-1}dt + \int_1^\infty e^{-t}t^{k-1}dt$$

로 구분하자.

(2-12)식의 뒷부분 $\int_1^\infty t^{k-1}e^{-t}dt$는 $k > 0$일 때 예비정리 2에 의해서

$$\begin{aligned}\lim_{t\to\infty} t^2 \cdot f(t) &= \lim_{t\to\infty} t^2 \left(e^{-t} \cdot t^{k-1}\right) \\ &= \lim_{t\to\infty} t^{k+1} \cdot e^{-t} \\ &= 0\end{aligned}$$

따라서 이 적분은 수렴한다.

(2-12)식의 앞부분 $\int_0^1 t^{k-1}e^{-t}dt$는

i) $k \geq 1$이면 피적분함수는 연속이니 따라서 적분가능하다.

ii) $0 < k < 1$일 때 만일 $\alpha + k - 1 > 0$ 즉 $1 - k < \alpha < 1$일 때는 예비정리 4에 의해서 $\lim_{t\to 0} t^\alpha \cdot f(t) = \lim_{t\to 0} t^\alpha \cdot (t^{k-1}e^{-t}) = 0$이므로 수렴한다.

한편 $k \leq 0$일 때 (2-11)식은 분명히 발산한다. 즉 Gamma 함수는 k가 양수인 값에 대해서만 정의된다는 것을 확인할 수 있다.

이제 일반적인 적분형태인 (2-5)식의 Gamma함수에 대해서

$$\mathcal{L}[t^k] = \int_o^\infty t^k e^{-st} dt = \frac{\Gamma(k+1)}{s^{k+1}}, \qquad (s > 0)$$

한편 (2-7)식에서 $\dfrac{\Gamma(k+1)}{s^{k+1}} = \dfrac{k}{s} \cdot \dfrac{\Gamma(k)}{s^k}$가 된다. 아니면 Gamma함수의 정의에 의해

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

를 부분적분법에 의해 적분함으로써

(2-13) $$\Gamma(k+1) = k\Gamma(k)$$

가 유도된다.

이제 k가 음수인 경우에 $\Gamma(k)$를 정의하기 위하여

(2-14) $$\Gamma(k) = \frac{\Gamma(k+1)}{k} = \frac{\Gamma(k+2)}{k(k+1)} = \cdots$$

으로 두자.

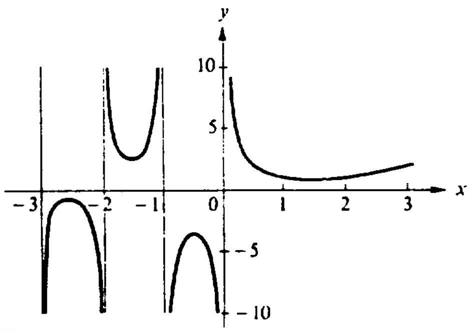
(2-13)식은

$$\lim_{k \to 0+} \Gamma(k) = \infty, \qquad \lim_{k \to 0-} \Gamma(k) = -\infty$$

이고 또 k가 양의정수일 때 $\Gamma(-k)$는 불연속이므로 $k = 0, -1, -2, \cdots$ 일 때는 의미가 없다. 그러나 $-1 < k < 0$인 임의의 값을 (2-13)의 좌변에 대입함으로써 $\Gamma(-k)$에

대한 값이 우변식으로 표시된다.

이러한 방법으로 점화식을 사용함으로써 $-1 < k < 0$일 때 $\Gamma(-\frac{1}{2}) = \Gamma(\frac{1}{2})/-\frac{1}{2}$, $\Gamma(-0.09) = \Gamma(0.1)/-0.9$, $\Gamma(k)$값은 물론 $-2 < k < 1$, $-3 < k < -2 \cdots$ 등의 모든 k값을 찾을 수 있고, 결국 0및 음의정수가 아닌 모든 k에 대한 $\Gamma(k)$를 계산할 수 있다. 즉, k가 음의정수가 아닐 때 $\Gamma(k)$는 (2-11)에 의한 이상적분의 값 및 점화식을 이용한 (2-13)에 의해 그림과 같이 도시되고, 이것은 또 $\Gamma(k+1) = k!$



그 림 : $y = \Gamma(x)$

에 의해 $\frac{1}{2}! = \Gamma(\frac{3}{2})$, $(-\frac{1}{2})! = \Gamma(\frac{1}{2}) \cdots$ 등 일반적인 계승의 값을 찾아낼 수 있다.

예 제 2. $\left(-\frac{1}{2}\right)!$을 계산하여 보자

풀 이 : $\frac{1}{2}! = \Gamma\left(\frac{3}{2}\right) = \int_0^\infty e^{-x} x^{\frac{1}{2}}\, dx$

$x = y^2$ 으로 두면

$$\frac{1}{2}! = 2\int_0^\infty e^{-y^2} y^2\, dy \quad 즉,\ \frac{1}{2}! = 2\int_0^\infty e^{-z^2} z^2\, dz$$

따라서

$$\left(\frac{1}{2}!\right)^2 = 4\int_0^\infty e^{-z^2} z^2\, dz \int_0^\infty e^{-y^2} y^2\, dy$$

$$= 4\int_0^\infty \int_0^\infty e^{-(y^2+z^2)} y^2 z^2\, dy\, dz$$

이제 $z = r cos\theta, \quad y = r sin\theta$ 로 두면

$$\left(\frac{1}{2}!\right)^2 = 4\int_0^\infty dr \int_0^{\frac{\pi}{2}} r^5\, e^{-r^2}\, sin^2\theta\, cos^2\theta\, d\theta$$

그런데 $\int_0^{\frac{\pi}{2}} sin^2\theta\, cos^2\theta\, d\theta = \frac{\pi}{16}$, 부분적분법에 의해 $\int_0^\infty e^{-r^2} r^5\, dr = 1$

따라서 $\left(\frac{1}{2}!\right)^2 = \frac{\pi}{4}$ 즉, $\frac{1}{2}! = \frac{\sqrt{\pi}}{2}$ 이고 (2-13)의 점화식에 의해

$$\left(-\frac{1}{2}!\right) = \sqrt{\pi} \text{ 가 된다.}$$

3. Gamma 함수의 주요성질

Gamma함수의 중요한 성질로써

$$\frac{1}{\Gamma(k)} = k e^{rk} \lim_{n\to\infty}\left[(1+k)(1+\frac{k}{2})\cdots(1+\frac{k}{n})e^{-k-\frac{k}{2}-\cdots-\frac{k}{n}}\right]$$

여기서 r는 Euler-Mascherini상수로써

$$r = 1 + \sum_{n=2}^{\infty}\left(\frac{1}{n} + \log\frac{n-1}{n}\right) = 0.5772\cdots.$$

이제 $\Gamma(k) = k^{k-\frac{1}{2}}e^{-k}\sqrt{2\pi}e^{\theta(k)/12k)}, \quad k > 0$

(단, $\theta(k)$는$0 < \theta(k) < 1$ 이 되는 k의 함수)

로 표시하면

$$\lim_{k\to\infty}\frac{\Gamma(k+1)}{k^{k+\frac{1}{2}}\sqrt{2\pi}e^{-k}} = 1$$

k가 정수일 때 이것은 $k!$에 대한 Stirling 근사식 : $k! \sim k^{k+\frac{1}{2}}\sqrt{2\pi}e^{-k}$ 을 유도한다.

만일 $k = 10$이면 좌변 $k! = 3.629 \times 10^6$이고 우변은 3.599×10^6이 된다.

한편 제1종의 Euler 적분이라고 불리어진 (1-5)의 Beta함수

$$\beta(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$$

는 xy-평면 위상의 $Q = (0,\infty) \times (0,\infty)$ 상에서 정의되는 그 변수의 함수로써

$x < 1, \quad y < 1$일 때 이 적분은 이상적분이지만 Gamma 함수에서 보여지는 것과

유사한 방법으로 수렴함을 보일 수 있다.

이제 $t = \sin^2\theta$ 를 대입하면

$$\beta(x,y) = 2\int_0^{\frac{\pi}{2}} sin^{2x-1}\theta cos^{2y-1}\theta\,d\theta$$

한편 Gamma 함수에서 $t = u^2$을 대입하면 $\Gamma(x) = 2\int_0^{\infty} u^{2x-1}e^{-u^2}\,du$ 로 표시되

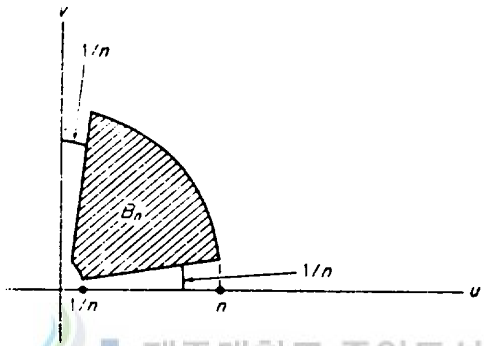
고

$$\Gamma(x)\Gamma(y) = 4\left(\int_0^{\infty} u^{2x-1}e^{-u^2}\,du\right)\left(\int_0^{\infty} v^{2x-1}e^{-v^2}\,dv\right)$$

$$\Gamma(x)\Gamma(y) = \int\int_Q e^{-u^2-v^2}u^{2x-1}v^{2y-1}\,du\,dv$$

단, $Q = (0,\infty)\times(0,\infty)$ 이고 $A_n = \left[\frac{1}{n}, n\right]\times\left[\frac{1}{n}, n\right]$ 이다.

이제 Q를 극좌표인 (γ,θ)로 치환하면



$$B_n = \{(r,\theta)\mid \frac{1}{n}\le r\le n,\quad \frac{1}{n}\le\theta\le\frac{\pi}{2}-\frac{1}{n}\}$$

이 되고, 따라서

$$\Gamma(x)\Gamma(y) = 4\lim_{n\to\infty}\int\int e^{-u^2-v^2}u^{2x-1}v^{2y-1}\,dudv$$

$$= 4\lim_{n\to\infty}\int_{\frac{1}{n}}^{\frac{\pi}{2}-\frac{1}{n}}\left(\int_{\frac{1}{n}}^{n} e^{-r^2}(rcos\theta)^{2x-1}(rsin\theta)^{2y-1}\,rdr\right)d\theta$$

$$= 4\lim_{n\to\infty}\left(\int_{\frac{1}{n}}^{\frac{\pi}{2}-\frac{1}{n}} cos^{2x-1}\theta\, sin^{2y-1}\theta\, d\theta\right)\left(\int_{\frac{1}{n}}^{n} r^{2(x+y)-1}e^{r^2}dr\right)$$

즉 $\Gamma(x)\Gamma(y) = \beta(x,y)\Gamma(x+y)$,

결국

$$\beta(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

이 유도된다.

Ⅲ. Gamma 함수의 응용

Gamma함수는 다양하게 응용되는데, 특히 확률과 통계에서 분포함수의 모형으로써, 또 추정과 여러가지 통계적 검정방법에서 이용되고 있다.

1. Gamma 분포

Gamma 함수 : $\Gamma(k) = \int_0^\infty e^{-t} t^{k-1} dt; k > 0$

에서 $t = \frac{x}{\theta}\,(\theta > 0)$로 두면

$$\Gamma(k) = \int_0^\infty \left(\frac{x}{\theta}\right)^{k-1} e^{\frac{x}{\theta}} \cdot \left(\frac{1}{\theta}\right) dx$$

양변을 $\Gamma(k)$로 나누면

$$1 = \int_0^\infty \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}} \quad x > 0$$

로써 확률밀도함수의 형태를 갖추게 된다.

<u>정 의 2 :</u> 확률밀도함수 X의 pdf가

$$f(x;\theta,k) = \begin{cases} \dfrac{1}{\theta^k \cdot \Gamma(k)} x^{k-1} e^{-\frac{x}{\theta}} & : x > 0 \\ = 0 & : x \leq 0 \end{cases}$$

(단, $k > 0$, $\theta > 0$.)

일 때 X는 Gamma분포를 갖는다고 한다.

Gamma분포는 특히 대기시간에 대한 확률모형으로서 활용되고 있다. 예를 들어 수명검사에 있어서 사망시까지의 대기시간 또는 어떤 특정 물체의 고장시까지의 시간 등은 Gamma분포를 갖는 확률변수이다. 이제 이를 확인하기 위하여 Poisson 과정의 공리를 가정하고 길이 ω인 구간을 時區間이라고 하자. 특히 확률변수 W를 k가 정해진 양의정수일 때 사망등으로 k개의 변화를 얻는 데 걸리는 시간이라고 하자.

W의 누적분포함수

$$G(w) = P(W \leq w) = 1 - P(W > w)$$

이다. $w > 0$에 대하여 사건 $W > w$는 길이 w인 시구간에 있어서 k변화보다 적게 존재하는 사건과 동치이다. 즉 확률변수 X가 길이 w 인 구간에서의 변화의 수라고 하면

$$P(W > w) = \sum_{x=0}^{k-1} P(X = x) = \sum_{x=0}^{k-1} \frac{(\lambda\omega)^x e^{-\lambda\omega}}{x!}$$

이고

$$\int_{\lambda\omega}^{\infty} \frac{z^{k-1}e^{-z}}{(k-1)!}dz = \sum_{x=0}^{k-1} \frac{\lambda\omega^x e^{-\lambda\omega}}{x!}$$

이 성립된다. 왜냐하면

$$\int_u^{\infty} \frac{1}{\Gamma(k)} z^{k-1}e^{-z}\,dz = \frac{1}{(k-1)!}\{z^{k-1}(-e^{-z})|_{\mu}^{\infty} + \int_{\mu}^{\infty}(k-1)z^{k-2}e^{-z}\}dz$$

$$= \frac{1}{(k-1)!}\mu^{k-1}e^{-\mu} + \frac{1}{k-2)!}\int_{\mu}^{\infty} z^{k-2}e^{-z}dz$$

$$= \frac{1}{(k-1)!}\mu^{k-1}e^{-\mu} + \frac{1}{(k-2)!}\mu^{k-2}e^{-\mu}$$

$$+ \frac{1}{(k-3)!}\int_{\mu}^{\infty} z^{k-3}e^{-z}dz$$

$$= \cdots\cdots$$

$$= \cdots\cdots$$

$$= \frac{\mu^{k-1}e^{-\mu}}{(k-1)!} + \frac{\mu^{k-2}e^{-\mu}}{(k-2)!} + \cdots + \frac{\mu^2 e^{-\mu}}{2!} + \frac{\mu e^{-\mu}}{1!} + e^{-\mu}$$

$$= \sum_{x=0}^{k-1} \frac{\mu^x e^{\mu}}{x!}$$

따라서 $\omega > 0$에 대해서 누적분포함수

$$G(\omega) = 1 - \int_{\lambda\omega}^{\infty} \frac{z^{k-1}e^{-\mu}}{\Gamma(k)}dz$$

$$= \int_0^{\lambda\omega} \frac{z^{k-1}e^{-z}}{\Gamma(k)} dz$$

이고, $\omega \le 0$ 에 대해서 $G(\omega) = 0$ 이다.

이제 $z = \lambda y$로 치환하면

$$G(\omega) = \begin{cases} \int_0^{\omega} \frac{\lambda^k y^{k-1} e^{-\lambda y}}{\Gamma(k)} & : \omega > 0 \\ 0 & : \omega \le 0 \end{cases}$$

이다. 따라서 W의 pdf는

$$g(\omega) = G'(\omega) = \begin{cases} \frac{\lambda^k \omega^{k-1} e^{-\lambda\omega}}{\Gamma(k)} & : \omega > 0 \\ 0 & : \omega \le 0. \end{cases}$$

즉 W는 $k = k$이고 $\theta = \frac{1}{\lambda}$ 인 Gamma분포를 갖는다는 것이 확인된다.

한편 W 가 최초의 변화까지의 대기시간 즉 $k = 1$이면

$$g(\omega) = \begin{cases} \lambda e^{-\lambda\omega} & : \omega > 0 \\ 0 & : \omega \le 0 \end{cases}$$

이고 이때 W 는 지수분포(Exponential distribution)를 갖는다고 한다.

2. x^2분포

Gamma분포는 연속형인 여러가지 비의 확률변수에 대한 모형이 된다. 예를 들어 소득의 분포는 두 모수 k와 θ를 적절히 취함으로써 Gamma분포에 의해 모형화될 수 있다. 즉, Gamma분포에서 $\theta = 2,\ k = \dfrac{\nu}{2}$가 되는 특수한 경우를 생각하자.

$$f(x, 2, \frac{\nu}{2}) = \begin{cases} \dfrac{1}{\Gamma(\frac{\nu}{2}) \cdot 2^{\frac{\nu}{2}}} x^{\frac{\nu}{2}-1} \cdot e^{-\frac{x}{2}} & : x > 0 \\ 0 & : x \le 0 \end{cases}$$

을 갖는 연속형 확률변수 X는 자유도 ν인 χ^2 분포를 갖는다고 하고 $\chi^2(\nu)$로 표기된다. 즉 $X \sim GAM(2, \dfrac{\nu}{2})$이면 $X \sim \chi^2(\nu)$이다.

<u>정 리 1.</u> $Y \sim \chi^2(\nu)$이면 $E(Y) = \nu,\ V(Y) = 2\nu$이다.

<u>정 리 2.</u> $X \sim GAM(\theta, k)$이면

$$Y = \frac{2X}{\theta} \sim \chi^2(2k).$$

왜냐하면 Y의 적률생성함수를 $M_Y(t)$라고 하면

$${}_Y(t) = M_{\frac{2X}{\theta}}(t) = M_X\left(\frac{2t}{\theta}\right) = (1-2t)^{-\frac{2t}{2}} \sim \chi^2(2k)$$

따라서 $X \sim GAM(\theta, k)$이고 F를 Gamma 분포함수, 그리고 자유도 k인 카이제곱 분포함수를 $H(y; \nu)$라고 하면 $F_Y(x) = H(2x/\theta\,;\, 2k)$로 표시된다.

예 제 3. 어떤 성분이 고장나기까지의 시간(年)은 $\theta = 3$이고 $k = 2$인 Gamma분포에 따른다. 그 성분의 90%가 고장나지 않고 작동할 보증기간은?

풀 이 : $P(X \le \chi_{0.1}) = H(2\chi_{0.1}/\theta\,;\,2k) = 0.1$

$$\frac{2\chi_{0.1}}{\theta} = \chi^2_{0.1}(2k)$$

따라서 $\chi_{0.1} = \dfrac{\theta \cdot \chi_{0.1}(2k)}{2}$

한편 $\theta = 3,\ k = 2$일 때

$\chi_{0.1} = \dfrac{3 \cdot \chi_{0.1}(4)}{2} = \dfrac{32 \times 1.06}{2} = 1.59$ 년

<u>정 리 3.</u> $Z \sim N(0,1)$이면 $Z^2 \sim \chi^2(1)$ 이다.

<u>정 리 4.</u> $X_1, X_2, \cdots, X_n$이 $N(\mu, \sigma^2)$으로부터의 확률분포일 때

$$\sum_{i=1}^{n} \frac{(X_i - \nu)^2}{\sigma^2} \sim \chi^2(n)$$

$$\frac{n(X - \mu)^2}{\sigma^2} \sim \chi^2(1)$$

한편 분산의 분포로 추정량 $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - X)^2$으로 둘 때 다음을 만족한다.

i) $\bar{X}$ 와 $X_i - \bar{X}$: $i = 1, 2, \cdots, n$ 은 독립이다.

ii) $\bar{X}$와 S^2은 서로 독립이다.

iii) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

위의 정리들의 증명은 생략한다.

예 제 4. X를 한 회사에서 생산하는 밧데리(Battery)의 수명(단위는 개월)이라고 하자. 이제 $X \sim N(60, 36)$이라고 가정하자. 만일 $S^2 \le 54.63$이면 $\sigma^2 = 36$을 기각하고 $S^2 < 54.63$이면 기각하지 않는다고 한다. $\sigma^2 = 36$인 주장을 기각하게 될 확률 즉

$$P(S^2 \ge 54.63) = P\left(\frac{24S^2}{36} \le 36.42\right)$$

$$= 1 - H(36.42; 24) = 0.05$$

3. χ^2검정

χ^2 검정은 1900년에 Karl Pearson에 의해서 제안된 χ^2 분포의 대표적인 응용 사례이다. 확률변수 X_i를 $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \cdots, n$이라 하고 $X_1, X_2, \cdots, X_n$을 서로 확률적으로 독립이라 하자.

이를 변수의 결합 pdf는

$$\frac{1}{\sigma_1\sigma_2\cdots\sigma_n(2\pi)^{n/2}} exp\left[-\frac{1}{2}\sum_1^n\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right]; \quad -\infty < x_i < \infty$$

이다. 여기서 지수에 의해 정의된 $\sum \frac{(X_i - \mu_i)^2}{\sigma_i^2} \sim \chi^2(n)$ 임을 앞의 정리에서 확인할 수 있다. 즉 이 사실이 χ^2검정의 수학적 기초가 된다. 이제 근사 χ^2분포를 갖는 몇가지 확률변수를 생각하자. X_1을 $B(n, p_1)$ 이라고 하자.

$Y = \frac{X_1 - np_1}{\sqrt{np_1(1-p_1)}}$은 $n \to \infty$일 때 N(0, 1)인 극한분포를 가지므로

$Z = Y^2 \sim \chi^2(1)$임을 보이자. Y 의 분포함수를 $G_n(y)$ 라 하면

$$\lim_{n\to\infty} G_n(y) = N(y) \ ; \quad -\infty < y < \infty$$

여기서 $N(y)$는 $N(0,1)$ 의 분포함수이다.

$H_n(z)$는 각 양의정수 n 에 대해

$Z = Y^2$의 분포함수라고 하면 $z \leq 0$ 일 때

$$\begin{aligned} H_n(Z) &= P(Z \leq z) = P(-\sqrt{z} \leq Y \leq \sqrt{z}) \\ &= Gn(\sqrt{z}) - Gn(-\sqrt{z}) \end{aligned}$$

한편 $N(y)$는 연속이니

$$\begin{aligned} \lim_{n\to\infty} H_n(z) &= N(\sqrt{z}) - N(-\sqrt{z}) \\ &= 2\int_0^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} d\omega \end{aligned}$$

여기서 $\omega^2 = \nu$로 치환하면

$$\lim_{n \to \infty} H_n(z) = \int_0^z \frac{1}{\Gamma(\frac{1}{2})2^{\frac{1}{2}}} \nu^{\frac{1}{2}-1} e^{-\frac{\nu}{2}} d\nu$$

즉, 이것은 $\chi^2(1)$ 이다.

X_1을 $B(n, P_1)$에서 $X_2 = n - X_1$, $p_2 = 1 - p_1$ 이라고 하자. Y^2을 Z대신 Q_1으로 표시하면

$$Q_1 = \frac{(X_1 - np_1)^2}{np_1(1-p_1)}$$

$$= \frac{(X_1 - np_1)^2}{np_1} + \frac{(X_1 - np_1)^2}{n(1-p_1)}$$

$$= \frac{(X_1 - np_1)^2}{np_1} + \frac{(X_2 - np_2)^2}{np_2}$$

일반적으로

$$Q_{k-1} = \sum_{k=1}^{k} \frac{(X_i - np_i)^2}{np_1}$$

여기서 $n \to \infty$일 때 Q_{k-1}는 $\chi^2(k-1)$인 극한분포를 가짐이 증명된다.

이제 $X_1, X_2, \cdots, X_n$의 결합 pdf가 모수 $n, p_1, \cdots, p_{k-1}$을 갖는 다항 pdf 이고 가설 H_0 ; $p_1 = p_{10}$, $p_2 = p_{20}$, $\cdots$, $p_{k-1} = p_{k-1,0}(p_k = p_{ko} = 1 - p_{10} - \cdots - p_{k-1,0})$ 을 생각하자. 만일 가설 H_0가 참이면 확률변수

$$Q_{k-1} = \sum^{k} \left[\frac{(X_1 - np_{io})^2}{np_{io}} \right]$$

은 자유도 $k-1$인 근사 χ^2 분포를 갖는다. H_0가 참일 때 np_{io}는 X_i의 기대치인 때문에 Q_{k-1}의 관찰치는 H_0가 참이면 너무 커서는 안된다는 것을 알 수 있다. 이제 자유도 $k-1$하에서 $P(Q_{k-1} \geq c) = \alpha$가 되는 c를 구할 수 있다. 이때 α는 이 검정의 유의수준이다. 만일 가설 H_0가 Q_{k-1}의 관찰치는 적어도 크기가 c와 같을 때 기각되면 H_0의 검정은 근사적으로 α와 같은 유의수준을 갖는다.

예 제 5. 주사위를 던져 하나의 양의정수를 선택한다.

즉 $A_i = \{x; x = i\}, \quad i = 1, 2, \cdots, 6$ 이다.

가설 $H_0 : P(A_i) = p_{i0} = \dfrac{1}{6} \quad (i = 1, 2, \cdots, 6)$

이 모든 경우에 대해 근사 5%의 유의수준에서 검정하고, 검정을 위해 같은 조건하에서 60회 득점 반복시행 되었다고 하자.

즉 $k = 6,\ np_{io} = 60 \times \dfrac{1}{6} = 10,\ i = 1, 2, \cdots, 6$이다. $Q_5 = \displaystyle\sum_{1}^{6} \frac{(X_i - 10)^2}{10}$ 이라 하자. H_0가 참이고 자유도 $k-1 = 6-1 = 5$일 때 $P(Q_5 \leq 11.1) = 0.05$이다.

만일 $A_1, A_2, \cdots A_6$의 실험도수가 13, 19, 11, 8, 5 및 4라고 가정하면 Q_5의 관찰되는

$$\frac{(13-10)^2}{10} + \frac{(19-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(5-10)^2}{10} + \frac{(4-10)^2}{10}$$

$$= 15.6$$

한편 $15.6 > 11.1$ 이니 가설 $P(A) = \frac{1}{6}(i = 1, 2, \cdots, 6)$은 유의수준 5% 에서 기각된

다.

예 제 6 .   두가지 형태의 완두콩의 교차속성에 관하여 생각하자.

멘델이론의 분류 a) 둥글고 황색, b) 주름지고 황색, c) 둥글고 녹색, d)주름지고 녹색

의 확률은 각각 $\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}$이다. 만일 160회 독립관찰에서 이들 각 분류의

관찰도수가 86, 35, 26, 13이라면 이들 데이타는 멘델이론과 일치하는가? 유의수준

1% 로 검정하여라.

풀 이 :

$$P_{10}(\text{둥글고, 노랑}) = \frac{6}{16}$$

$$P_{20}(\text{주름지고, 노랑}) = \frac{3}{16}$$

$$P_{30}(\text{둥글고, 노랑}) = \frac{3}{16}$$

$$P_{40}(\text{주름지고, 녹색}) = \frac{1}{10}$$

$$H_0\,;\,P_i = P_{io} \quad ; \quad i = 1,2,3,4.$$

여기서 H_0가 참이면

$$Q_3 = \sum_1^4 \frac{(X_1 - np_{io})^2}{np_{io}} \sim \chi^2(3)$$

$$= \frac{(36-90)^2}{90} + \frac{(35-30)^2}{30} + \frac{(26-30)^2}{30} + \frac{(13-10)^2}{10}$$

$$= \frac{16}{90} + \frac{25}{30} + \frac{16}{30} + \frac{9}{10}$$

$$= \frac{220}{90} = 2.44$$

한편 $\nu = 4$일 때 유의수준 1% 에서의 기각역 $Q_3 > 11.3$

따라서 $2.44 < 11.3$이니 H_0가 채택된다. 즉 이 자료는 멘델이론이 적합하다.

한편 확률실험의 결과가 두 속성으로 분류된다고 하자. 이들 각 속성 A, B는 각각 $A_1, A_2, \cdots, A_a$중의 오직하나이고 서로 배반이며 마찬가지로 B도 $B_1, B_2, \cdots, B_b$ 중 하나이면서 서로 배반적이다. $P_{ij} = P(A_i \cap B_j)$, $i = 1,2,\cdots,a$, $j = 1,2,\cdots,b$ 라고 하자. 확률실험을 독립적으로 n회 반복하고, X_{ij}는 사건 $A_i \cap B_j$의 돗수를 표시한다. $A_i \cap B_j$와 같은 사건 $k = ab$가 존재하므로 확률변수

$$Q_{ab-1} = \sum_{j=1}^{b} \sum_{i=1}^{a} \frac{(X_{ij} - np_{ij})^2}{cp_{ij}}$$

은 $n \to \infty$일 때 자유도 $ab-1$인 근사적 χ^2 분포를 갖는다.

이제 속성 A와 속성 B의 독립성을 검정한다고 가정하자.

즉, 가설 H_0 ; $P(A_i \cap B_j) = P(A_i) \cdot P(B_j),\ i = 1, 2 \cdots, a,\ j = 1, 2, \cdots, b$ 만일

$P(A_i) = p_i,\ P(B_j) = p_j$ 라고 쓰면 $P_i = \sum_{j=1}^{b} p_{ij},\ P_{ij} = \sum_{i=1}^{a} p_{ij}$ 이고

따라서 $\sum_{j=1}^{b}\sum_{i=1}^{a} p_{ij} = \sum_{j=1}^{b} p_{ij} = \sum_{i=1}^{a} p_i = 1$ 이므로 가설 H_0 ; $P_{ij} = Pi\cdot, P_j$,

$i = 1, 2, \cdots, a,\ \ j = 1, 2, \cdots, b$로 수식화할 수 있다. H_0를 검정하기 위하여 P_{ij}를

$P_i,\ P_j$대치하여 Q_{ab-1}을 사용할 수 있다. 만일 P_{i0}와 P_{oj}가 미지이면 이 미지모수를

$$P_{io} = \frac{X_{io}}{n}, \quad X_{io} = \sum_{j=1}^{b} X_{ij}$$

$$P_{io} = \frac{X_{oj}}{n}, \quad X_{oj} = \sum_{i=1}^{a} X_{ij}$$

로 추정한다. 결국 확률변수

$$\sum_{j=1}^{b}\sum_{i=1}^{a} \frac{[X_{ij} - n(X_{io}/n)(X_{oj}/n)]^2}{n(X_{io}/n)(X_{oj}/n)}$$

은 H_0가 참이면 자유도가 $ab - 1 - (a + b - 2) = (a-1)\cdot(b-1)$인 근사 χ^2 분포를 갖는다.

예 제 7.　두 속성이면서 각각은 배반적으로 분류되는 형태의 확률실험을 200회 독립시행하니 다음 데이타를 얻었다.

	B_1	B_2	B_3	B_4
A_1	10	21	15	6
A_2	11	27	21	13
A_3	6	19	27	24

유의수준 5%로 A와 B의 속성의 독립성을 검정하라.

풀 이 :

$$P(A_i \cap B_j) = P(A_i) \cdot P(b_j), \ i = 1,2,3 \ \ j = 1,2,3,4.$$

자유도 $11 - 5 = 6$. 따라서

$$Q_6 = \sum_{j=1}^{4}\sum_{i=1}^{3} \frac{\left[X_{ij} - n(\frac{X_{io}}{n})(\frac{X_{0j}}{n})\right]^2}{n(\frac{X_{io}}{n})(\frac{X_{0j}}{n})} = 12.94$$

한편 유의수준 5%에서 $\chi^2(6)$의 기각역 $Q_6 > 12.6$이고 따라서 $12.94 > 12.6$이므로 H_0를 기각한다. 유의수준 5%에서 속성 A와 속성 B는 독립이 아니다.

참 고 문 헌

[1] Bain, L.J. Introduction to Prob. and Mathematical Statistics 1987, PWS publishers

[2] Craig, A and Hogg, R.V. Introduction to Math. Statistics 4^{th}ed, 1978.

[3] Kaplan, W. Advanced Calculus , 1984.

[4] Sokolnikoff, I.S Advanced Calculus; McGraw–Hill Book company, Inc, 1939.

[5] 劉東善外 理工系 統計學, 1985. 集賢社.

〈Abstract〉

Gamma Function and the Study of the Application

Pu, Suk-Hee

Mathematics Education Major
Graduate School of Education
Cheju National University
Cheju, Korea

Supervised by professor Kim, Ik-Chan

This paper begins with the historical consideration of Gamma function and researches the convergence and the major qualities of Gamma function by generalizing factorial n. Gamma distribution and χ^2-distribution are brought out in Gamma Function. The practical examples are shown which are actually being applied in the authorization of uniformity, independence and the good of fit applied in the theory of presumption authorization. to emphasize the importance of Gamma function is the purpose of this paper.

* This thesis submitted to the Committed of the Graduate School of Education, Cheju National University in partial fulfillment of the requirements for the degree of Master of Edcuation in June, 1992.

감 사 의 글

이 논문이 완성되기까지 연구에 바쁘신 가운데도 세밀한 지도를 하여주신 김익찬교수님께 감사드리며, 그동안 지도를 아끼지 않으신 수학교육과 수학과 교수님께 심심한 감사를 드립니다.

그리고 오늘에 이르기까지 따뜻한 마음으로 보살펴주신 부모님께도 또한 감사를 드립니다.

1992년 6월 일

부 숙 희