응용통계연구 제 9권 2호
1996년 9월 pp. 83-94

윷의 확률 추정에 대하여

박 진 경[1], 박 홍 선[2]

요 약

윷의 확률에 대한 기하학적 접근의 선행연구가 있었고, 이 논문은 그 선행연구를 보완, 발전시키는데 목적을 두고 있다. 특히, 윷의 확률이 '논리적'으로 계산되기 어려운 상황과 그 이유를 설명하고, 통계학적인 추정방법들을 제시하고 있다. 시중에 판매되고 있는 여러 가지 종류의 윷을 사용함으로써 실질적인 회귀선을 유도하였고, 이와 같이 추정된 확률을 통하여 도, 개, 걸, 윷, 모 의 출현빈도순서 및 예상확률을 구할 수 있게 되었다. 이러한 통계학적인 접근의 결과는 기초확률시간이나 기초통계학시간에 활용될 수 있을 것이다.

1. 서론: 선행연구의 요약

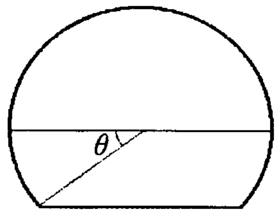
윷의 확률에 대한 '논리적' 연구는 김미경, 허명회(1995)에 의해 처음으로 시도되었다. 선행연구는 윷을 던질 때와 굴릴 때를 구분하였고, 각 경우에 윷을 한꺼번에 던질 때와 하나씩 던질 때로 구분하여 실험하였다. 그들은 먼저, 윷의 단면에서 측정되는 각도 θ (그림 1 참조)를 사용하여 '곡면이 출현할 확률' p'을 다음 식으로 계산하였다

$$p' = \tan^{-1}\left(\cos\theta / (\sin\theta + \frac{2\cos^3\theta}{3(\frac{\pi}{2} + \theta + \sin\theta\cos\theta)})\right) / \pi$$

선행연구에 사용된 윷의 각도는 17.6 ° 이고 이에따른 '평면이 출현할 확률'은 $p = 1 - p'$ =0.6713 이었다. 이때 윷을 한 개씩 던졌을 때의 추정값은 $\hat{p} = 0.664$ 이었고 4개를 한꺼번에 던졌을 때 추정값은 $\hat{p} = 0.579$ 이었다.

---

1) 경기도 용인시 모현면 왕산리 89, 한국외국어대학교 통계학과 석사과정.
2) 경기도 용인시 모현면 왕산리 89, 한국외국어대학교 통계학과 조교수.



<그림 1> 윷의 단면

2. 최우추정법과 베이즈추정법

본 연구는 우리가 즐기고 있는 윷놀이에서는 윷을 굴리거나 윷을 하나씩 던지는 경우가 없다는 사실을 바탕으로, 윷 4개를 한꺼번에 던질 경우에 대한 윷의 확률에 대하여 알아보기로 한다. 이 경우, 선행연구에서 살펴본 바와 같이, '논리적 확률', 0.6713 은 실제 '경험적 확률', 0.579 와 차이를 보이게 되었는데 그 이유를 다음과 같이 설명할 수 있겠다.

첫째, 윷이 떨어지는 순간, 충격에너지를 바닥이 완전히 흡수한다는 가정과 윷가락이 바닥과 수평 하게 떨어진다는 가정은 실제상황과는 거리가 멀고, 이러한 가정과 현실의 차이는 윷을 한꺼번에 던질 경우에, 윷위에 윷이 떨어지는 상황이 발생하여 더욱 커지게 된다.

둘째, 윷 위에 윷이 떨어지는 순간, 윷의 구르는 효과가 가중되어서 그 만큼 평면이 출현할 가능성이 줄어들게 된다.

위와 같은 실제상황을 고려하면, 윷을 4개 한꺼번에 던질 경우 '평면이 출현할 확률'을 논리적으로 구하는 것은 거의 불가능하다고 판단되어진다. 그러나, 4개의 윷을 한꺼번에 던질 경우에, 윷과 윷이 부딪히는 상황과, 윷위에 윷이 떨어지는 상황, 혹은 아무런 충돌 없이 떨어지는 상황등은 모든 윷가락에게 동일하게 적용되고 또 각 상황의 확률이 일정하다(던지는 패턴이 일정할 경우)는 가정을 하면, 임의의 윷 1개의 '평면이 출현할 확률'은 4개의 윷에 대해 서로 동일하다고 가정할 수 있다. 또한, 하나의 윷의 출현이 '평면'이 되는 사건은 다른 윷의 출현과는 독립이라고 볼 수 있게 된다. 그렇다면, 비록 논리적으로 그 확률을 계산할 수는 없지만, 윷놀이가 이항분포 $B(4, p)$를 따른다는 가정은 무리 없이 할 수 있게 된다.

1) 최우추정량: $\hat{p}_{MLE}$

만일 n번 윷을 던졌다면, 이는 $4n$회의 독립적인 베르누이 시행이므로,

$$\hat{p}_{MLE}=\frac{\text{도의 도수}\times 1+\text{개의 도수}\times 2+\text{걸의 도수}\times 3+\text{윷의 도수}\times 4}{4n}$$

가 된다. 이때 성공의 횟수는 '평면이 출현한 윷가락 수'를 의미한다.

2) 베이즈추정량: $\hat{p}_{Bayes}$

어떤 임의의 종류의 윷이 '평면이 나올 확률' p를 가지고 있고, 이는 Beta (α, β)라는 사전분포에서 출현한 확률변수라고 가정한다면, 이항우도함수에 따르는 짝분포(Conjugate Posterior)는 모수가 $\alpha+x$, $4n-x+\beta$ 인 베타분포임을 알 수 있다(Fruend and Walpole, 1980). 따라서, 확률 p에 대한 베이즈 추정량은

$$\hat{p}_{Bayes}=E(p|x)=\frac{\alpha+x}{\alpha+\beta+4n}$$

가 된다. 이때, α와 β 를 베타분포의 평균과 분산을 사용한 method of moment 추정법을 이용하여 구하면,

$$\hat{\alpha}=\bar{p}\left(\frac{\bar{p}(1-\bar{p})}{s_p^2}-1\right), \qquad \hat{\beta}=(1-\bar{p})\left(\frac{\bar{p}(1-\bar{p})}{s_p^2}-1\right)$$

가 계산된다. 이때, $\bar{p}$는 여러종류의 윷을 바닥의 상태를 달리한 상태에서 던져서 구한 12개의 확률의 근사치들에 대한 평균 0.5446 이고, s_p^2 는 그 확률들의 분산 0.0034 이다. 이를 사용하면, $\hat{\alpha}=39.18$, $\hat{\beta}=32.76$ 이 된다. 이 연구에서는 최우추정량과의 동등한 비교를 위하여 해당되는 상황의 윷을 제외시킨, 나머지 값들의 평균과 분산을 이용하였지만 그 결과는 기존의 평균과 분산을 사용하였을때와 거의 일치하고 있다. 세번째로 생각할 수 있는 추정량은, 사전분포가 균일분포(uniform distribution)이라는 가정 하에서 $\alpha=1$, $\beta=1$ 을 사용하는 것이다.

우리는 먼저 윷의 확률에 대한 추정을 언급하기 전에, 윷의 확률이 '윷의 종류'와 '바닥의 상태'에 따라 얼마만큼 영향을 받는지 살펴보았다. 다음은 실험에 사용된 4종류의 윷의 사양이다.

<표 1> 실험에 사용된 윷

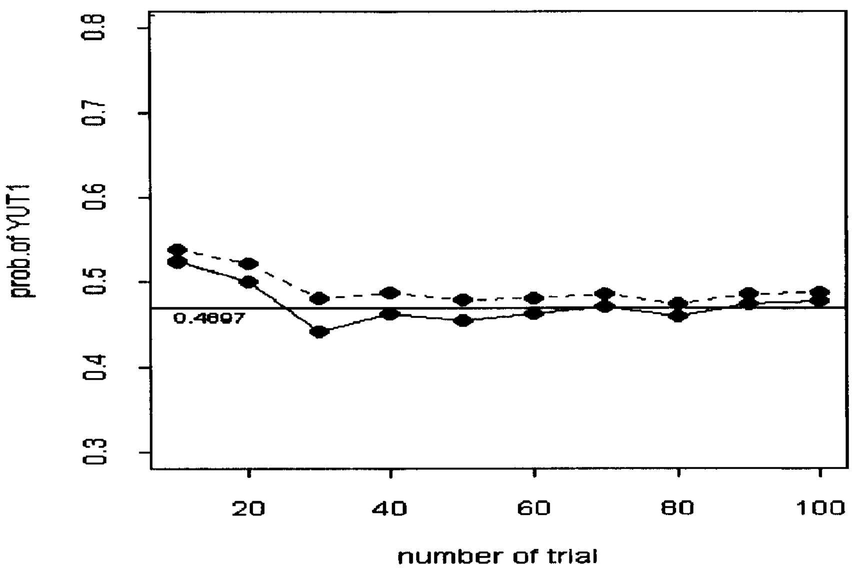
윷의 종류	1	2	3	4
길이(cm)	5.2	9.9	17.2	20.1
둘레(cm)	3.6	6.5	6.8	7.5
무게(g/개)	3.6	12.8	24.4	60.9
재질	플라스틱	나무	나무	나무
θ (°)	12.58	36.85	33.83	42.19

여기서, 각도 θ 의 측정은 3절에서 설명하기로 하고 실험에 사용된 바닥은 얇은 바닥(0.6 cm), 중간바닥(1.2 cm), 두꺼운 바닥(2.4 cm)이 사용되었다. 이 실험은 '바닥'과 '윷의 종류'에 의한 3x4 요인배치법이며, 각 처리조합에서는 100번씩 윷던지는 실험을 총 8회 실시함으로써 8개의 p 가 구해지게 되고, 따라서 우리는 '바닥효과'와 '윷 종류' 그리고 '바닥과 윷의 교호작용'에 의한 분산분석을 할 수 있게 되었다. (여기서 윷과 바닥은 변량인자로 간주하였다.)

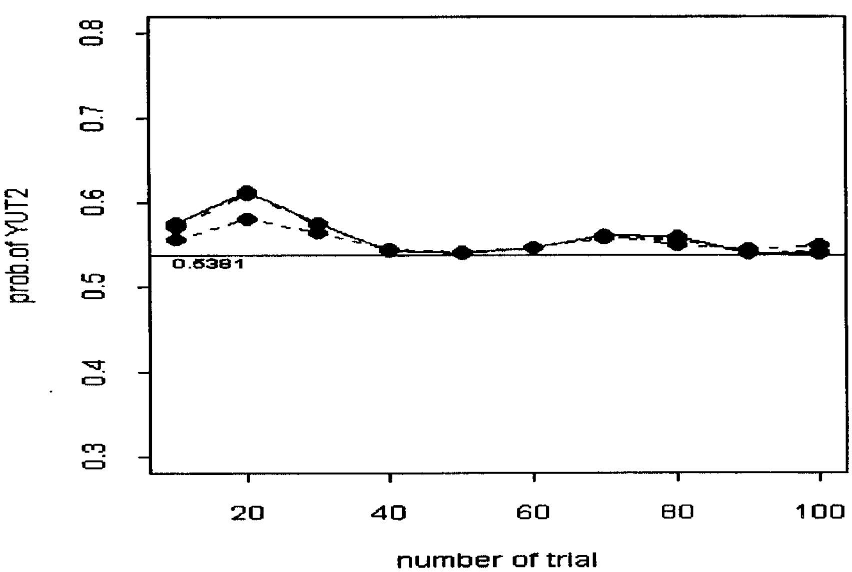
<표 2> 실험에 의한 분산분석

요인	자유도	제곱합	평균제곱	F
윷	3	0.27435833	0.09145278 (ϕ_A)	44.6911** (ϕ_A/ϕ_{AB})
바닥	2	0.01113451	0.00556725 (ϕ_B)	2.7206 (ϕ_B/ϕ_{AB})
윷*바닥	6	0.01227799	0.00204633 (ϕ_{AB})	2.9005* (ϕ_{AB}/ϕ_E)
오차	84	0.05926250	0.00070551 (ϕ_E)	
전체	95	0.35703333		

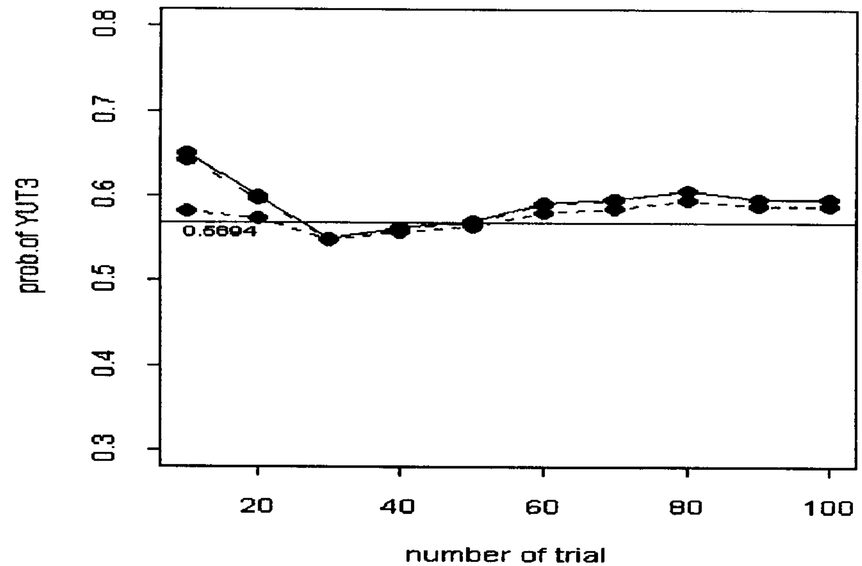
이 결과로써, 우리는 '윷의 종류'가 윷의 확률에 많은 영향을 주고 있음을 알 수 있고, 바닥 자체만으로는 유효한 영향을 주지 못하고 있음을 알 수 있다. 물론, 바닥의 차이를 극대화 시킬 경우에 유의한 차이를 유도할 수 있겠지만, 본 연구가 실제 벌어지는 윷놀이의 상황에 초점을 맞추었음을 다시 한번 주지하는 바이다. 아래의 <그림 2>-<그림 5>는 중간바닥을 사용했을 경우에 4종류의 윷에 대한 각각의 p의 근사치와, 시행횟수를 증가시킴에 따른 점근 추세이다. 실선은 $\hat{p}_{MLE}$를 의미하고, 점선은 베타사전을 이용한 $\hat{p}_{Bayes}$이며, 나머지 긴 점선은 균일 사전을 이용한 $\hat{p}_{Bayes}$이고 이는 $\hat{p}_{MLE}$와 거의 같이하고 있다.



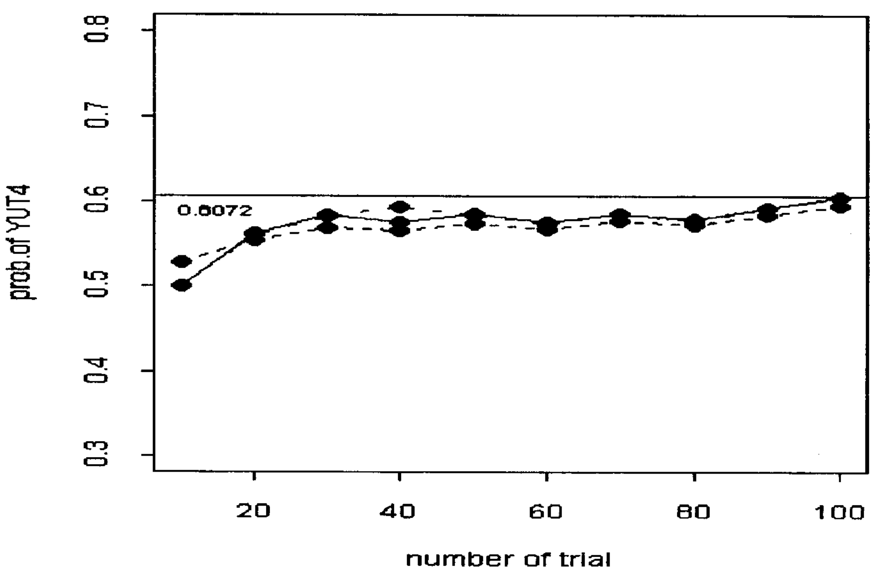
<그림 2> 윷 1에 대한 확률 추이



<그림 3> 윷 2에 대한 확률 추이



<그림 4> 윷 3에 대한 확률 추이



<그림 5> 윷 4에 대한 확률 추이

이제까지 살펴본 바와 같이, 세개의 추정량은 모두 실행횟수가 증가함에 따라 대응하는 p의 근사치에 수렴하고 있으며, 특히 40번 이상에서는 매우 근사한 모습을 보이고 있다. 따라서, 어느 추정량을 사용하더라도, 40번 정도 던진 값은 매우 신뢰성 있는 값이라고 할 수 있겠다. 또한, 세개의 추정량 가운데에서 베타사전분포를 이용한 베이즈 추정량이 비교적 안정적(stable)인 점근추세를 보이고 있고, 이는 특히 n이 작을 때 더욱 두드러졌다. 결론적으로 말하면, 임의의 사람이 임의의 윷을 가지고, 임의의 바닥에서 던질 경우, '평면이 출현할 확률'을 구하려면, 기존의 평균과 분산 등을 사용한 베타사전분포의 베이즈방법을 사용하길 추천하되 n이 40이상일 경우는 어느 추정량을 사용하여도 결과는 비슷함을 알 수 있었다.

3. 회귀선을 사용한 추정

2절에서는 윷의 확률을 몇 번의 시험(trials)을 통하여 그 확률을 추정하는 방법들을 논하였다. 그러나, 윷을 던지지 않고 그 값을 추정할 수는 없을까? 만일 논리적 확률이 실재확률과 일치한다면 단면의 각도를 측정하여(혹은 계산하여) 얻겠지만, 앞서 언급한 대로 논리적 계산방법은 윷을 4개 한꺼번에 던졌을 때는 사용할 수 없게 된다. 이미 우리는 윷의 확률이 '윷의 종류'에 따라 차이가 나고 있음을 보았으므로, 이 단원에서는 윷의 확률을 윷의 길이, 무게, 각도 등을 이용하여 회귀선을 구하여 보았다. 먼저 시중에서 판매되고 있는 서로 다른 12개의 종류의 윷과 한종류의 윷을 한꺼번에 40회 던져서 구한 총 160회 시행의 $\hat{p}_{MLE}$ 은 <표 3>과 같다.

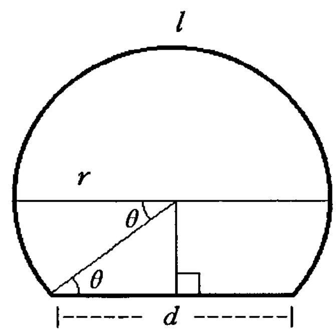
<표 3> 윷의 종류와 40번 던진 결과

윷종류	길이(cm)	각도(˚)	무게(g)	확률
1	5.2	12.58	3.6	0.444
2	9.9	36.85	12.8	0.544
3	17.2	33.83	24.4	0.569
4	20.1	42.19	60.9	0.581
5	10.5	31.81	18.0	0.531
6	10.0	37.85	16.3	0.631
7	17.7	41.41	46.7	0.550
8	17.5	27.63	32.3	0.594
9	17.0	30.74	24.5	0.544
10	17.0	29.62	23.8	0.556
11	17.0	32.84	24.3	0.563
12	18.5	37.45	47.5	0.600

여기서, 윷의 단면의 각도를 구하는데는 다음과 같은 수식을 사용하였다. 윷의 바닥 면의 길이를 d, 나머지 곡면의 둘레를 l, 반지름을 r이라 할 때 (그림 6 참조), 윷의 d 와 l 을 측정한 후에,

$$\frac{d}{2} = r\cos\theta\ , \quad 2\pi : 2\pi r = \pi + 2\theta : l$$

의 연립방정식을 풀면 이는 $f(\theta) = 2l\cos\theta - 2d\theta - d\pi = 0$ 의 해가 되며, 이렇게 구한 θ 에 $180/\pi$를 곱해주면 된다.

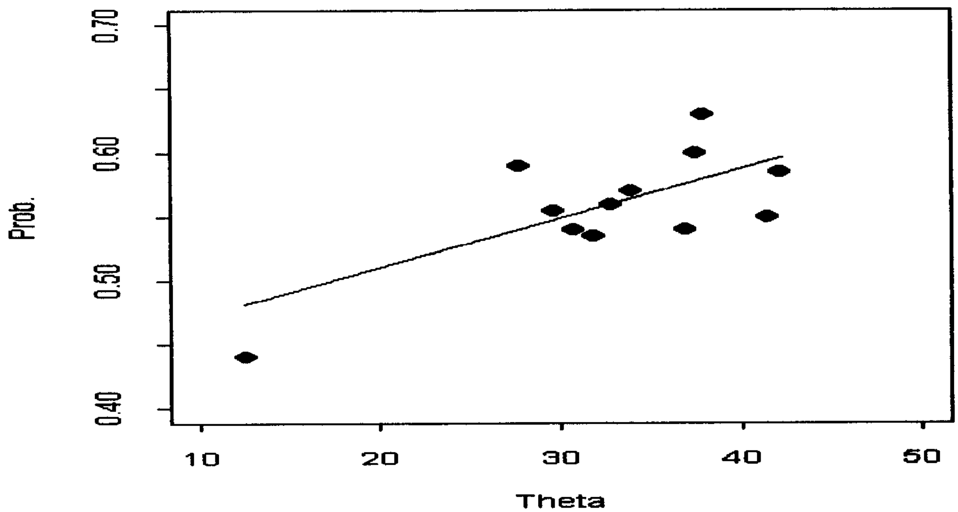


<그림 6> 윷의 각도 측정

이 논문에서는 FORTRAN과 IMSL을 이용하여 $f(\theta) = 0$ 의 해를 구하였다. 이와 같이 구한 각도(θ)와 무게(weight), 그리고 길이(length)에 의한 '윷의 확률'의 회귀식을 구하기 위하여 교호작용을 포함한 이차 다항식을 고려하여서 stepwise로 변수를 선택하였다. 그 결과를 요약하면, 추정회귀식은 각도 θ 에 의존하는 일차형식을 나타내고 있다.

$$\hat{p} = 0.41358 + 0.004387 * \theta$$

이를 그림으로 나타내면 <그림 7>과 같다. 이 회귀선을 구함에 있어서 각도가 20도 이하인 실제 판매되는 윷을 구하기가 어려웠고 오직 1개의 윷 밖에 구할 수 없었다. 하지만, 각도가 커짐에 따라서 윷의 평면이 출현할 확률이 커질 것이라는 직관적인 예상과 일치되는 면을 고려할 때, 의미 있는 회귀선이라고 할 수 있다.



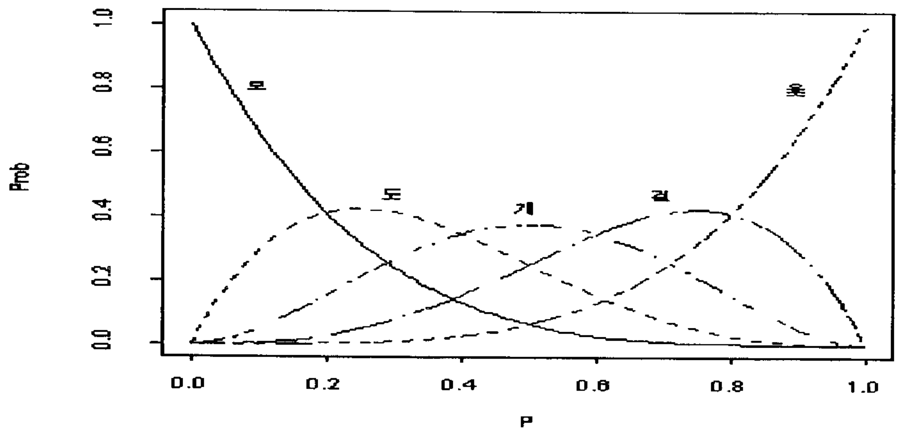
<그림 7> 윷의 각도에 따른 근사 확률

4. 윷의 확률의 응용

윷놀이를 할 때, 임의의 윷을 가지고 시작하기 전에 40번 정도 던저서 그 확률을 추정할 수도 있겠지만, 그만큼의 시간적 여유가 없을 때는 적은 횟수의 시행을 통한 베이즈 추정법을 사용하던지 혹은 앞에서 구한 회귀선을 이용하여 그 평면이 출현될 확률을 추정하는 것이 바람직할 것이다. 그 추정방법이 어떤 것이든, 일단 $\hat{p}$가 주어진다면, 이제 우리의 관심은 도, 개, 걸, 윷, 모 의 출현 빈도순일 것이다. 이항분포 $B(4, \hat{p})$를 사용하여 그 확률을 구하면 $\binom{4}{x}\hat{p}^{x}(1-\hat{p})^{(1-x)}$, $x=0, 1, 2, 3, 4$ 를 이용하여 다음과 같은 그림을 얻을 수 있다.

<표 4> 구간에 따른 출현 빈도순서

확률구간	출현확률이 큰 순서
$0.4 < p < 0.5$	개>도>걸>모>윷
$0.5 < p < 0.6$	개>걸>도>윷>모
$0.6 < p < 0.6135$	걸>개>도>윷>모
$0.6135 < p < 0.7101$	걸>개>윷>도>모



<그림 8> '평면이 출현할 확률' 에 따른 도, 개, 걸, 윷, 모의 확률

<표 3>에서 알 수 있듯이, 시중에서 판매되는 대부분의 윷은 p가 0.5와 0.6 사이에 존재하고 있고 이는 <표 4>와 <그림 8>에서 알 수 있듯이 '개'의 출현이 '걸'의 출현보다 빈번하게 나타난다. 그러나, 윷의 모양이 곡면이 많은 형태로 발전할수록 p의 값은 0.6에 근접하게 되고, 나가서는 '걸'의 빈도수가 '개'의 빈도수 보다 많은 형태의 윷으로 판정될 것이다.

5. 맺음말

우리가 즐겨하는 윷놀이는 윷 4개를 한꺼번에 던지는 게임이고, 이 경우에, 논리적 확률은 적지 않은 문제점을 안고 있다. 이러한 상황에서 우리는 경험적 추정법인 최우추정량과 베이즈 추정량을 통하여 윷의 확률을 구해 보았고, 그 결과, 40번 이상의 횟수에서는 두 추정량이 서로 근사하게 나타나고 있음을 보여주었다. 그 외에 윷의 확률에 영향을 주는 많은 요소를 생각하여 보았는데, 그 중에 각도만이 유의하게 영향을 주는 요소로 분석되었다. 또한, 윷의 단면의 각도를 그 곡면의 둘레와 평면의 너비를 측정하여서 FORTRAN 프로그램을 이용하여 계산한다면, 우리가 앞서 구한 회귀선을 사용하여 임의의 윷을 미리 던져보지 않고도 그 확률을 추정할 수 있는 방법을 제시해 보았다. 이와 같이 윷의 확률을 바르게 추정하고, 제시된 도표 <그림 8>를 사용하여 우리는 도, 개, 걸, 윷, 모 의 출현빈도순서 및 그에 해당하는 확률을 예측할 수 있게 된다. 이와 같은 윷의 확률에 대한 정보는 '말을 업을 것인가' 혹은 '앞으로 나갈 것인가' 등의 의사결정에 결정적 역할을 하게 될 것이다.

참고문헌

[1] 김미경, 허명회 (1995). 윷의 확률, Proceedings of the Spring Conference, Korean Statistcal Society, 91-96.

[2] Johnson, N. L. & Kotz, S. (1970). *Continuous Univariate Distributions-2,* John Wiley Sons, New York, 37-53.

[3] Freund, J. E. & Walpole, R. E. (1980). *Mathematical Statistics,* Prentice Hall, New Jersey, 204.

감사의 글

저자들은 이 논문의 수정에 도움을 주신 익명의 두분 심사위원들께 감사를 드립니다.

On Estimation of the Probability of Yut

Jinkyung Park[3], Heungsun Park[4]

Abstract

The probability of Yut was calculated by using the physical property in previous study, but this article suggested empirical estimators for probability of Yut. In practice, physics-based probability imposes too strong assumptions, which result in the difference between the calculated probabilies and empirical relative frequencies. Experiment shows the probabilities of Yut depend on the integrated shape of Yut rather than the floor type. Maximum likelihood estimator and empirical Bayes estimators are compared and all turn out to be almost identical for more than 40 trials. For smaller number of trials, Bayes estimators are recommended for its stability. Regression approach is also adopted as an easy-to-use method without empirical trials.

3) Dept. of Statistics, Hankuk University of Foreign Studies, Yongin.
4) Assistant Professor, Dept. of Statistics, Hankuk University of Foreign Studies, Yongin.