The Korean Journal of Applied Statistics (2016) DOI: http://dx.doi.org/10.5351/KJAS.2016.29.7.1421
29(7), 1421–1428

Analysis of freshness of rice depending on packing materials using MANOVA

Seong-Ju Kim[a,1]

[a]School of Business, Sungkyunkwan University

(Received October 20, 2016; Revised November 7, 2016; Accepted November 14, 2016)

Abstract

This paper concerns the freshness of rice depending on packing materials using MANOVA. Freshness of rice is measured in terms of moisture content and rice flavor. Ordinary paper and charcoal-coated paper are compared as packing materials. Storing places are considered as a block. The bivariate observations of moisture content and the rice flavor are compared using MANOVA for a completely randomized block design. It is observed that there is a significant difference between ordinary paper and charcoal-coated paper. Therefore we apply ANOVA for moisture content and rice flavor, respectively. Significant differences are observed for the moisture content but not for the rice flavor.

Keywords: completely randomized block design, MANOVA

1. 서론

국내 농업에서 쌀은 가장 큰 비중을 차지하고 생존과 직결되는 매우 중요한 작물이다. 우리나라는 경지면적이 좁고 경지 정리가 잘 되어 있지 않기 때문에 아무리 기계화를 한다고 해도 규모의 경제에 의한 효율성은 떨어질 수밖에 없다. 한편 미국의 경우는 대규모 영농이 가능하고 중국의 경우는 인건비마저도 저렴하기 때문에 우리 쌀의 가격 경쟁력은 현저히 떨어질 수밖에 없다.

그런데 국내 쌀 생산은 연이은 풍년인데 비해 1인당 쌀 소비량은 계속 감소하고 있다. The Korean Statistical Information Service (2016)에 의하면 Table 1.1과 같이 1인당 연간 쌀 소비량은 1980년에 132.4 kg이던 것이 2015년에는 62.9 kg으로 반 이하로 줄어들었다. 그 결과 현재 정부의 쌀 비축량은 175만 톤에 달하며 올해 추수가 끝나면 비축량은 200만 톤을 넘을 것이라고 한다. 설상가상으로 정부 쌀 비축량 중에서 50만 톤은 수입쌀이며 이는 각종 자유무역협정(FTA)으로 인하여 우리나라가 의무적으로 수입한 물량이다.

Lee와 Lee (2016)에 의하면 국내 쌀 시장을 보호하기 위해 지난 20년 동안 미뤄왔던 관세화 유예 조치를 해제하고 쌀 시장을 전면 개방하게 되면 공급과잉으로 가격하락을 초래하고, 농가수지가 악화될 것으로 예상하고 있다. 그렇다고 해서 저렴한 수입쌀을 제쳐두고 우리 쌀을 먹자고 애국심에 호소하는 데는 한계가 있다. 따라서 가격 경쟁력을 갖춘 수입쌀과 경쟁하려면 우리 쌀의 품질을 높이는 것이 하나의 대안이 될 수 있다.

[1]School of Business, Sungkyunkwan University, 25-2, Sungkyunkwan-Ro Jongro-Gu, Seoul 03063, Korea.
E-mail: seongju@skku.edu

Table 1.1. Average rice consumption per person by year (kg)

년도	1980년	1990년	2000년	2010년	2015년
1인당 연간 쌀 소비량	132.4	119.6	93.6	72.8	62.9

아무리 품질이 우수한 쌀을 생산하더라도 유통 과정에서 변질된다면 아무 소용이 없기 때문에 쌀의 신선도를 오래 보존할 수 있는 포장 재질을 개발할 필요가 있다. 따라서 본 논문에서는 다변량 및 단일변량 분산분석을 이용하여 포장 재질에 따라 쌀의 신선도가 어떻게 달라지는지 분석해 보고자 한다. 포장 재질로는 일반 종이와 종이 뒷면에 숯가루를 코팅한 종이(이하 숯 종이)를 비교하고자 한다.

쌀의 신선도는 쌀의 수분 함량과 식미치로 측정하고자 한다. 그 이유는 쌀의 신선도를 객관적으로 측정할 수 있는 측도로서 수분 함량과 식미치 말고는 다른 것을 생각하기 어렵기 때문이다. 여기서 쌀의 수분 함량은 15%일 때 가장 이상적이라고 한다. 쌀의 수분 함량이 15% 이상이면 도정하기 어려울 뿐만 아니라 쌀이 변질되기 쉽다고 하고 수분함량이 15% 미만이면 밥맛이 떨어진다고 한다. 그리고 식미치(食味置)는 밥맛이 얼마나 좋은지 100점 만점으로 측정한 값으로서 식미치가 높을수록 밥맛이 좋은 쌀로 평가하며 보통 70점 이상이면 밥맛이 괜찮은 쌀이라고 한다.

원래 밥맛이 얼마나 좋은지는 밥을 직접 지어서 숙련된 패널에 의해 평가하는 것이 원칙이지만 비용과 시간을 절약하기 위해 식미 측정기에 의해 평가하는 것이 보편화 되어 있다. 식미 측정기는 쌀의 성분(아밀로스 함량, 단백질 함량 등)을 측정하여 다중회귀식에 의해 밥맛을 추정해주는 기계이며 Min (1998)에서 자세히 설명하고 있다. 식미치는 식미 측정기 제조 회사에 따라 달라지는데 본 논문의 식미치는 일본 TOYO 회사의 식미 측정기로 측정한 값이다.

일반적으로 가정에서 쌀을 보관하는 장소로서 냉장고, 다용도실, 싱크대, 발코니를 생각해 볼 수 있다. 본 논문에서는 쌀을 보관하는 장소 각각을 블록(block)으로 간주하고 각 처리 조합에 대해 3번씩 반복 측정하였다. 그런데 쌀은 햇빛이 들지 않고, 통풍이 잘되고, 서늘한 곳에 보관해야 신선도를 오래 유지할 수 있다고 알려져 있다. 이러한 관점에서 보면 냉장고는 햇빛이 안 들고 서늘하지만 통풍 안 되는 단점이 있다. 다용도실은 햇빛이 덜 들고, 통풍은 잘 되지만 서늘하지 않다는 단점이 있다. 싱크대는 햇빛은 안 들지만 통풍이 안 되고 서늘하지 않다는 단점이 있다. 발코니는 통풍은 잘 되지만 햇빛이 잘 들고 서늘하지 않다는 단점이 있다.

2절에서는 본 논문에서 분석하려고 하는 데이터를 수집한 과정을 설명하고 그 데이터를 요약하고 표현해 보았다. 3절에서는 완전 확률화 블록 설계(completely randomized block design)에 의한 통계 분석에 대하여 설명하고 있다. 끝으로 4절에서는 본 논문의 결론과 연구의 한계에 대하여 언급하고 있다.

2. 실제 데이터

본 논문에서 분석하려고 하는 실제 데이터는 다음과 같이 수집되었다. 이러한 데이터를 수집하게 된 동기는 두 가지가 있다. 첫째, 숯이 항균 작용을 하고 보습 효과가 있다고 하는데 숯 종이로 쌀을 포장하면 일반 종이에 비해 쌀의 신선도를 향상시킬 수 있을지에 관심이 있었다. 둘째, MANOVA 모형을 만족시킨다고 볼 수 있는 실제 데이터가 매우 드물기 때문에 이 실험계획을 철저히 하면 MANOVA 모형을 적용시킬 수 있는 하나의 실제 사례가 될 수 있다고 생각하였다.

1. 처리(treatments): 포장 재질로는 일반 종이와 숯 종이 두 가지를 비교하였다.
2. 블록(block): 가정에서 쌀을 보관하는 장소로서 냉장고, 다용도실, 싱크대, 발코니를 고려하였으며 이를 블록으로 간주하였다.

Table 2.1. Raw data of moisture content and rice flavor by packing materials and storing places

	수분함량				식미치			
	냉장고	다용도실	싱크대	발코니	냉장고	다용도실	싱크대	발코니
일반종이	14.1	13.3	13.1	13.0	71	71	68	70
	14.0	13.2	13.1	13.0	72	74	70	70
	14.2	13.3	13.1	13.0	72	69	70	71
숯 종이	14.0	13.8	13.6	13.7	72	68	69	72
	14.1	13.8	13.4	13.5	72	70	71	72
	14.1	13.8	13.7	13.5	72	69	70	72

Table 2.2. Means of moisture content by packing materials and storing places

	수분함량				행 평균
	냉장고	다용도실	싱크대	발코니	
일반종이	14.10	13.27	13.10	13.00	13.37
숯 종이	14.07	13.80	13.57	13.57	13.75

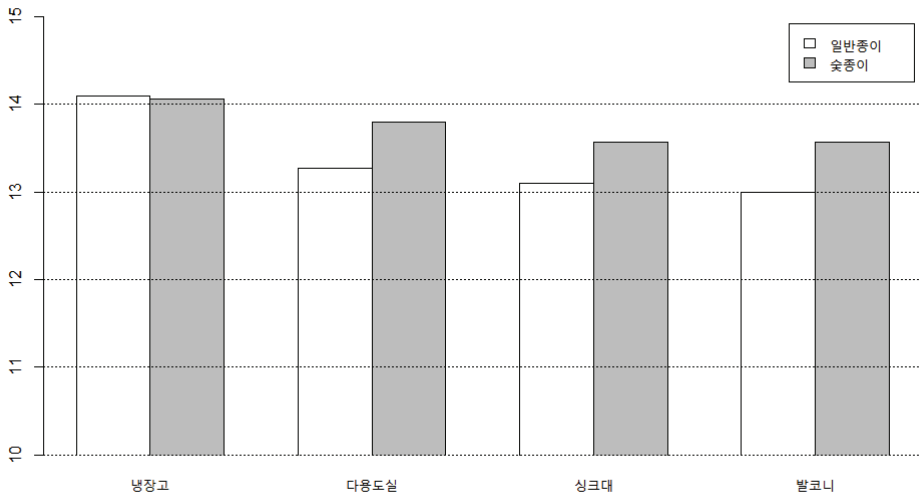

Figure 2.1. Bar chart of moisture content by packing materials and storing places.

3. 처리 조합(treatment combination): 이 실험에서 처리 조합은 모두 2(처리) × 4(블록) = 8개가 있으며 갓 도정한 쌀 16 kg을 2 kg 씩 담아서 8개 처리 조합을 만들었다.
4. 반복 횟수(replications): 각 처리 조합에 대하여 3번씩 반복 측정하였다.
5. 확률화(randomization): 확률화는 블록에 대해 실현하였다.
6. 데이터의 수집: 쌀을 보관하고 나서 일주일 후, 쌀의 신선도를 농업기술센터에서 측정하였다. 쌀의 신선도는 쌀의 수분 함량과 식미치로 측정하였으며 수분 함량의 단위는 %이고 식미치는 TOYO 식미 측정기로 측정한 점수(100점 만점)이다. 수분 함량과 식미치를 측정한 연구원은 어떤 처리, 어떤 블록에서 나온 샘플인지를 알 수 없도록 눈가림 시행(blind experiment)을 철저히 적용하였다.

이러한 과정을 거쳐 수집된 데이터를 Table 2.1에 나타내었다. 우선 수분 함량에 대하여 보관 장소 별

Table 2.3. Means of rice flavor by packing materials and storing places

	식미치				행 평균
	냉장고	다용도실	싱크대	발코니	
일반종이	71.67	71.33	69.33	70.73	70.67
숯 종이	72.00	69.00	70.00	72.00	70.75

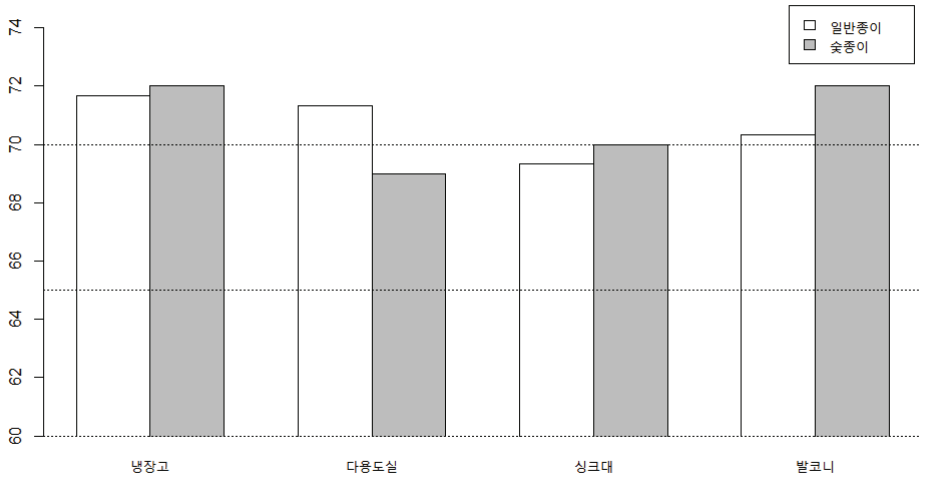

Figure 2.2. Bar chart of rice flavor by packing materials and storing places.

일반 종이와 숯 종이의 수분 함량의 평균을 Table 2.2에 나타내었다. Figure 2.1은 Table 2.2를 시각적으로 표현한 것이다. 여기서 단순통계량으로서 평균 이외의 표준편차도 구하려 했으나 0이 되는 경우가 많아서 생략하였다. 마찬가지로 식미치에 대하여 보관 장소 별 일반 종이와 숯 종이의 식미치의 평균을 Table 2.3에 나타내었다. Figure 2.2는 Table 2.3를 시각적으로 표현한 것이다. 수분 함량과 식미치의 상관계수는 일반 종이($n = 12$)인 경우 $r = 0.41$이고, 숯 종이($n = 12$)인 경우 $r = 0.12$이었다. 일반 종이와 숯 종이를 종합($n = 24$)하면 $r = 0.28$이었다.

주관적인 판단이지만 Figure 2.1에 의하면 수분 함량은 포장 재질에 따라 차이가 나는 것으로 보인다. 그러나 Figure 2.2에 의하면 식미치는 포장 재질에 따라 차이가 나지 않는 것으로 보인다.

3. 완전 확률화 블록 설계(completely randomized block design)에 의한 통계 분석

3.1. 다변량 분산분석(multivariate ANOVA; MANOVA)

3.1에서는 수분 함량과 식미치를 이변량 벡터라고 간주하고 다변량 분산분석을 적용하고자 한다. 즉, $Y, \mu, \alpha, \beta, \epsilon$를 이변량 벡터라고 간주하면 다변량 분산분석 모형은 식 (3.1)과 같다.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}, \quad i = 1, 2; \;\; j = 1, \ldots, 4; \;\; k = 1, \ldots, 3. \tag{3.1}$$

식 (3.1)에서,

Y_{ijk}는 i번째 처리($i = 1$은 일반 종이, $i = 2$는 숯 종이), j번째 블록($j = 1$은 냉장고, $j = 2$는 다

용도실, $j = 3$은 싱크대, $j = 4$는 발코니) 그리고 k번째 반복에서 관측한 수분 함량과 식미치를 나타내는 이변량 벡터이고,

α_i는 i번째 처리 효과로서 $\sum_{i=1}^{2} \alpha_i = 0$,

β_j는 j번째 블록 효과로서 $\sum_{j=1}^{4} \beta_j = 0$,

$\alpha\beta_{ij}$는 α_i와 β_j의 교호작용 효과로서 $\sum_{i=1}^{2} \alpha\beta_{ij} = 0$, $\sum_{j=1}^{4} \alpha\beta_{ij} = 0$,

ϵ_{ijk}는 i번째 처리, j번째 블록, 그리고 k번째 반복에서 발생한 오차이며 이변량 정규분포를 따른다.

식 (3.1)에서 블록 효과 β_j와 교호작용 효과 $\alpha\beta_{ij}$를 고려하지 않는다면 식 (3.1)은 'Multivariate Behrens-Fisher Problem'으로 귀결되며 Kim (1992)에 의해 연구된 바 있다. 본 논문에서는 오차항에 대해 정규성과 더불어 등분산성도 가정하고자 한다.

식 (3.1)에서 다음 세 가지 가설에 대하여 관심을 갖으며 Johnson과 Wichern (2012)에 의하면 세 가지 가설은 Wilks' Lambda가 작을 때 기각된다.

첫째, 포장 재질의 주효과가 없다는 가설 H_{b1} ; $\alpha_1 = \alpha_2 = 0$.

둘째, 블록 효과가 없다는 가설 H_{b2} ; $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$.

셋째, 교호작용 효과가 없다는 가설 H_{b3} ; $\alpha\beta_{11} = \alpha\beta_{12} = \cdots = \alpha\beta_{24} = 0$.

Table 2.1의 이변량 데이터에 대하여 식 (3.1)을 적용해 보자. 첫째, 포장 재질의 주효과가 없다는 가설 H_{b1}에 대한 Wilks' Lambda는 0.0974로서 p-값은 0.0001 보다 작았다. 둘째, 블록 효과가 없다는 가설 H_{b2}에 대한 Wilks' Lambda는 0.0246으로서 p-값은 0.0001 보다 작았다. 셋째, 교호작용 효과가 없다는 가설 H_{b3}에 대한 Wilks' Lambda는 0.1310으로서 p-값은 0.0001 보다 작았다.

오차에 대한 정규성을 검정하기 위하여 잔차를 이용하여 Shapiro-Wilk 검정 결과 $W = 0.9759$이고 p-값은 0.8102이므로 정규성을 기각할 수 없었다. Shapiro-Wilk 검정 이외의 다른 정규성 검정, 예컨대 Kolmogorov-Smirnov 검정이나 Anderson-Darling 검정을 수행하여도 p-값은 통상적인 유의수준 보다 높게 나타났기 때문에 정규성을 기각할 수 없었다.

따라서 우리는 식 (3.1)을 받아들일 수 있으며 쌀의 신선도 즉, 수분 함량과 식미치는 포장 재질에 따라 차이가 나고 그리고 보관 장소에 따라 차이가 난다는 당연한 사실을 확인할 수 있다.

3.2. 단일변량 분산분석(univariate ANOVA)

제 3.1장에서는 수분 함량과 식미치를 이변량 벡터라고 간주하고 다변량 분산분석을 적용하였다. 그 결과 포장 재질의 주효과가 없다는 가설은 기각되었다. 따라서 우리는 수분 함량에 차이가 있는지 또는 식미치에 차이가 있는지에 관심을 갖게 된다.

우선 수분 함량 Y에 대한 단일변량 분산분석 모형은 식 (3.2)와 같다.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}, \quad i = 1, 2; \;\; j = 1, \ldots, 4; \;\; k = 1, \ldots, 3. \tag{3.2}$$

식 (3.2)에서,

Y_{ijk}는 i번째 처리($i = 1$은 일반 종이, $i = 2$는 숯 종이), j번째 블록($j = 1$은 냉장고, $j = 2$는 다용도실, $j = 3$은 싱크대, $j = 4$는 발코니) 그리고 k번째 반복에서 관측한 수분 함량이고,

Table 3.1. ANOVA table for moisture content

	d.f.	SS	MS	F	p-value
요인	1	0.8817	0.8817	132.25	< 0.0001
블록	3	2.4150	0.8050	120.75	< 0.0001
교호작용	3	0.3550	0.1183	17.75	< 0.0001
오차	16	0.1067	0.0067		
총합	23	3.7584			

Table 3.2. ANOVA table for rice flavor

	d.f.	SS	MS	F	p-value
요인	1	0.0417	0.0417	0.03	0.8597
블록	3	17.1250	5.7083	4.42	0.0191
교호작용	3	13.1250	4.3750	3.39	0.0441
오차	16	20.6667	1.2917		
총합	23	50.9584			

α_i는 i번째 처리 효과로서 $\sum_{i=1}^{2} \alpha_i = 0$,

β_j는 j번째 블록 효과로서 $\sum_{j=1}^{4} \beta_j = 0$,

$\alpha\beta_{ij}$는 α_i와 β_j의 교호작용 효과로서 $\sum_{i=1}^{2} \alpha\beta_{ij} = 0$, $\sum_{j=1}^{4} \alpha\beta_{ij} = 0$,

ϵ_{ijk}는 i번째 처리, j번째 블록, 그리고 k번째 반복에서 발생한 오차이며 $\epsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$이다.

식 (3.2)에서 우리는 다음 세 가지 가설에 대하여 관심을 갖는다.

첫째, 포장 재질의 주효과가 없다는 가설 H_{u1} ; $\alpha_1 = \alpha_2 = 0$.

둘째, 블록 효과가 없다는 가설 H_{u2} ; $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$.

셋째, 교호작용 효과가 없다는 가설 H_{u3} ; $\alpha\beta_{11} = \alpha\beta_{12} = \cdots = \alpha\beta_{24} = 0$.

Table 2.1의 수분 함량에 대하여 식 (3.2)를 적용하면 Table 3.1의 분산분석표를 얻게 된다. Table 3.1에 의하면 가설 H_{u1}, 가설 H_{u2}, 가설 H_{u3}에 대한 p-값은 모두 0.0001 보다 작기 때문에 세 가지 가설은 통상 적용되는 어떠한 유의수준에서도 기각된다. 교호작용에 대해 살펴보자. 블록1(냉장고)에서 숯 종이는 일반 종이에 비해 수분 함량의 차이가 없는 것으로 보인다. 그러나 블록2(다용도실), 블록3(싱크대), 블록4(발코니)에서는 숯 종이가 일반 종이에 비해 수분 함량이 평균적으로 높게 나타났다. 이는 납득할 수 있는 결과라고 생각된다. 왜냐하면 숯은 일반적으로 보습력이 강하다고 알려져 있지만 냉장고에서는 그 효과를 발휘하기 어렵기 때문이다.

앞에서 쌀의 신선도를 측정하는 첫 번째 기준인 수분 함량에 대하여 단일변량 분산분석을 한 것과 똑같이 두 번째 기준인 식미치에 대하여 단일변량 분산분석을 적용하면 Table 3.2의 분산분석표를 얻게 된다. Table 3.2에 의하면 통상 적용되는 유의수준에서 포장 재질의 주효과가 없다는 가설 H_{u1}은 기각되지 않지만 블록 효과가 없다는 가설 H_{u2}와 교호작용 효과가 없다는 가설 H_{u3}는 기각된다. 교호작용에 대해 살펴보자. 블록1(냉장고)과 블록3(싱크대)에서는 숯 종이는 일반 종이에 비해 식미치에 차이가 없는 것으로 보인다. 그러나 블록2(다용도실)에서는 일반 종이가 숯 종이에 비해 식미치가 평균적으로 높게 나타났고 반대로 블록4(발코니)에서는 숯 종이가 일반 종이에 비해 식미치가 평균적으로 높게 나타났다. 이러한 현상에 대해서는 좀 더 탐구해 볼 필요가 있다고 생각한다.

4. 결론 및 연구의 한계

본 논문에서는 포장 재질에 따른 쌀의 신선도를 다변량 분산분석과 단일변량 분산분석을 이용하여 분석하였다. 쌀의 신선도는 수분 함량과 식미치로 측정하였다. 수분 함량과 식미치를 이변량 벡터라고 간주하고 다변량 분산분석을 수행하였다. 그 결과 포장 재질(일반 종이와 숯 종이)과 쌀을 보관하는 장소(냉장고, 다용도실, 싱크대, 발코니)에 따라 수분 함량과 식미치 벡터는 유의한 차이를 보였다. 따라서 수분함량과 식미치를 각각 단일변량으로 간주하고 단일변량 분산분석을 수행하였다. 그 결과 수분 함량에서는 포장 재질과 쌀을 보관하는 장소에 따라 유의한 차이를 보였다. 그러나 식미치에서는 쌀을 보관하는 장소에 따라 유의한 차이를 보였지만 포장 재질에 따라 유의한 차이를 발견할 수 없었다. 이 결론은 2절에서 실제 데이터를 시각적으로 표현한 것과 일치하는 것을 알 수 있다.

본 논문에서는 포장 재질로 일반 종이와 숯 종이 두 가지만 고려하였고 포장 용량을 2 kg 한 가지만 고려하였다. 또한 실험 기간이 일주일이라서 초단기 변화만 관측하였다. 이러한 점은 본 논문의 한계이며 앞으로 더 폭 넓은 연구를 수행할 필요가 있다고 생각한다.

References

Johnson, R. A. and Wichern, D. W. (2012). *Applied Multivariate Statistical Analysis* (6th ed), Pearson, London.

Kim, S. J. (1992). A practical solution to the multivariate Behrens-Fisher problem, *Biometrika*, **79**, 171–176.

Lee, H. R. and Lee, H. S. (2016). Impacts of delayed tariffication on rice in Korean farm's income, *Journal of International Trade and Industry Studies*, **21**, 1–30.

Min, B. G. (1998). Recent trends of evaluating the quality of rice, *Applied Biological Chemistry*, **5**, 4–10.

The Korean Statistical Information Service (2016). *2015 grain consumption Survey*.

다변량 분산분석을 이용한 포장 재질에 따른 쌀의 신선도 분석

김성주[a,1]

[a]성균관대학교 경영대학

(2016년 10월 20일 접수, 2016년 11월 7일 수정, 2016년 11월 14일 채택)

요 약

본 논문에서는 포장 재질에 따라 쌀의 신선도가 어떻게 달라지는지 알아보기 위해 다변량 분산분석을 적용해 보았다. 쌀의 신선도는 쌀의 수분 함량과 식미치로 측정하였다. 포장 재질은 일반 종이와 종이 뒷면에 숯을 코팅한 종이를 비교하였다. 쌀을 보관하는 장소를 블록으로 간주하였다. 쌀의 수분 함량과 식미치를 나타내는 이변량 관측값에 대하여 완전 확률화 블록 설계에 의한 다변량 분산분석을 적용해 본 결과 포장 재질에 따라 유의한 차이를 보였다. 따라서 쌀의 수분 함량과 식미치 각각을 단일변량 관측값으로 간주하고 분산분석을 적용해 보았다. 그 결과 쌀의 수분 함량에 대해서는 포장 재질에 따라 유의한 차이를 보였으나 식미치에 대해서 유의한 차이 관측할 수 없었다.

주요용어: 다변량 분산분석, 완전 확률화 블록 설계, 윌크 람다

[1](03063) 서울특별시 종로구 성균관로 25-2, 성균관대학교 경영대학. E-mail: seongju@skku.edu