

패널요원 수행능력 평가에 사용된 분산분석, 상관분석, 주성분분석 결과의 비교

김상숙 · 홍성희 · 민봉기 · 신명곤

한국식품개발연구원 쌀이용연구센터

Evaluation of Panel Performance by Analysis of Variance, Correlation Analysis and Principal Component Analysis

Sang Sook Kim, Sung Hie Hong, Bong Kee Min and Myung Gon Shin

Rice Utilization Research Center, Korea Food Research Institute

Abstract

Performance of panelists trained for cooked rice quality was evaluated using analysis of variance, correlation analysis, and principal component analysis. Each method offered different information. Results showed that panleists with high F ratios ($p=0.05$) did not always have high correlation coefficient ($p=0.05$) with mean values pooled from whole panel. The results of analysis of variance for the panelists whose performance were extremely good or extremely poor were consistent with those of correlation analysis. Outliers designated by principal component analysis were different from the panelists whose performance was defined as extremely good or extremely poor by analysis of variance and correlation analysis. The results of principal component analysis discriminated the panelists with different scoring range more than different scoring trends depending on the treatments. Our study suggested combination of analysis of variance and correlation analysis provided valid basis for screening panelists.

Key words: panel evaluation, cooked rice, analysis of variance, correlation analysis, principal component analysis.

서 론

관능검사요원은 식품을 비롯한 일반 소비재의 관능적 특성 및 강도를 측정하는 유일한 도구로 사용되고 있으며 제품의 품질관리면에서 아주 중요한 역할을 한다. 이들 패널요원이 임무를 수행하는데 있어서의 성과는 패널요원의 선정 및 훈련에 사용된 엄격한 기준과 절차에 달려있다.

관능검사요원의 수행능력을 평가하는 방법에는 여러 가지가 있는데, Cross 등⁽¹⁾은 각 패널개인의 평가자료를 분산분석하여 F값의 순위에 의해 관능검사요원의 수행 능력을 평가하였으며, Gacula와 Singh⁽²⁾, Malek 등⁽³⁾, McDaniel 등⁽⁴⁾은 상관분석에 의해 패널요원을 평가하여 선정하였다. 최근 다변량분석은 관능검사분야에서 검사 항목이 많은 경우 자료의 양을 줄이고 문제들을 분류하는데 이론적이고 실제적인 관심의 대상이 되어왔다^(5,6). 주성분분석은 다변량자료의 한 분석기법으로 자료의 요

약과 새로운 해석을 그 목적으로 하고 있다⁽⁷⁾. 이 방법은 관능검사 평가항목이 많은 경우, 패널훈련동안 유용하게 사용될 수 있는데, Wu⁽⁸⁾ 등은 주성분 분석결과로 그래프를 그려 대부분의 패널과 다르게 시료를 평가하는 패널요원을 찾아낼 수 있다고 하였다. Powers⁽⁹⁾는 주성분분석과 상관분석은 같은 목적으로 사용될 수 있으며 모든 상관계수를 하나하나 계산하는 것보다는, 주성분분석이 더 낫다고 하였다. 패널수행능력 평가에는 이 외에도 군집분석⁽¹⁰⁾, control chart 방법⁽¹¹⁾ 등 여러가지 방법들이 제시되고 있다. 그러나 아직 수행능력을 평가하는 가장 신뢰성 있는 방법들에 관한 견해는 일치하지 않고 있으며, 특히 가장 보편적으로 사용되고 있는 분산분석, 상관분석, 그리고 최근 주목을 받고 있는 주성분분석 세가지 방법을 실제 예를 들어 비교한 문헌은 없다.

현재 우리나라에서의 관능평가는 여러분야에서 그의 중요성이 인식되기 시작했으나, 패널요원의 훈련 및 평가를 주도할 수 있는 전문관능요원이 부족한 현실이며 신뢰성 있는 관능검사결과를 얻기 위해 필수적인 검사목적에 맞는 패널요원의 선정 및 훈련은 그 중요성을 인식하면서도 세부적인 패널평가절차에 대한 연구를 찾아

Corresponding author: Sang Sook Kim, Rice Utilization Research Center, Korea Food Research Institute, Baek-hyun-dong san 46-1, Bundaang-gu, Seungnam-si, Kyunggi-do 463-420, Korea

보기 힘들다. 본 연구의 목적은 쌀밥의 품질평가요원의 검사수행능력을 평가하는데 이용되는 세 가지의 분석방법 즉, 분산분석, 상관분석, 주성분분석에 의한 결과를 비교하고, 이를 평가방법의 장단점과 더불어 가장 신뢰성있는 패널평가방법을 제시하는데 있다.

재료 및 방법

시료제조

쌀은 92년도산 낙동벼로 도정후 사용전까지 나이론/폴리에틸렌 복합 비닐백에 밀봉하여 저온저장(4°C) 하였다. 패널요원 평가에 사용된 시료는 쌀 900g을 수압 세미기(PR-7J, Aiho)로 동일조건에서 세미를 한 후 취반한 시료로서 표준시료외에 세 시료를 제시하였다. 표준시료는 쌀(수분함량 13.0%)에 1.4배의 가수율을 적용하여 물을 첨가한 후 1시간 침지시킨 다음 전기밥솥(MHC-510, 마마전기)을 이용하여 취반하였다. 첫번째 시료는 표준시료와 동일한 시료를 제시(숨겨진 표준시료)하였으며, 두번째 시료는 표준시료와 동일한 조건으로 취반하여 4시간동안 보온($72 \pm 2^{\circ}\text{C}$)하고 15시간 동안 냉장보관(4°C) 후 마이크로웨이브 오븐(ER-900MHB, 금성사)을 이용하여 딥힌(1kg의 밥을 600w 고주파 출력으로 5분간 가열) 시료를 사용하였다. 세번째 시료는 1.3배의 가수율을 적용하여 취반한 밥을 사용하였다.

시료 제시

취반 후 밥시료는 예비실험을 통해 확립해 놓은 표준방법을 이용하여 밥을 솔에서 다른 용기로 옮겨 담아 10분간 식힌 후 사기용기(8.5 cm × 5 cm, 지름 × 높이)에 50g을 담아 뚜껑을 덮어 제시하였다. 제시된 시료의 온도는 $27 \pm 2^{\circ}\text{C}$ 이었다. 각 패널요원에게 한번에 4개의 시료가 제시되었는데, 1개는 S라고 표시된 표준시료였으며 3개는 세자리 무작위 숫자가 표시된 평가시료였다. 시료 제시순서는 S로 표시된 표준시료는 맨 왼쪽에 제시했으며 나머지 세시료는 무작위로 배치하였다.

패널요원

패널평가에 참가한 패널요원은 한국식품개발연구원에 근무하는 건강한 남녀 40명으로서, 이들은 56명의 예비 패널요원으로부터 선정되어 2주간의 1차 훈련과정을 마친 사람들이었다. 예비패널로부터 패널선정은 ASTM (12)이 제시한 패널 선정방법을 일부 변형하여 실시하였는데, 2주간 소요되었으며, 기본맛 확인, 기본맛의 농도 순위정하기, 밥을 이용한 10번의 삼점검사를 포함하였다. 56명의 예비패널로부터 출석율과 평가성적에 의해 선정된 40명을 다시 2주간 훈련시킨 후 패널요원 평가과정에 사용하였다.

관능적 품질평가

쌀밥의 품질평가는 5개의 평가항목(향, 외관, 맛, 조

직감 및 종합적인 품질)으로 이루어졌으며, 평가방법은 표준시료와 비교하여 평가하는 13항목척도(1=표준보다 많이 낫다, 3=표준보다 보통 낫다, 5=표준보다 약간 낫다, 7=표준과 같다, 9=표준보다 약간 높다, 11=표준보다 보통 높다, 13=표준보다 많이 높다)를 사용하였다.

통계분석

40명의 패널요원에 의한 3시료의 관능적 품질평가는 4번 반복하였다. SAS⁽¹³⁾를 이용하여 시료간의 차이가 있는지 여부를 결정하기 위해 분산분석을 하였다. 전체적인 분산분석 후, 시료간의 차이가 있는 특성항목을 선별하여 그 항목들만을 이용하여 패널 수행능력 평가에 이용하였다. 분산분석, 상관분석, 주성분분석을 실시하였으며, 각 분석방법에 의한 패널요원의 수행능력을 비교하였다.

결과 및 고찰

실험의 전체 분산분석 결과 모든 특성항목에서 유의적인 차이가 있었기 때문에 모든 특성항목(향, 외관, 맛, 조직감 및 종합적 품질평가)별로 모든 패널요원에 대해 분산분석을 실시하였으며, 전체 패널의 평균값과 각 패널요원의 평균평가점수간의 상관계수를 구하였다. 먼저 세 가지 분석방법을 효과적으로 비교하기 위해 40명의 패널로 부터 거의 모든 경우의 상황을 보여줄 수 있는 12명의 패널을 선정하였다. 선정된 12명의 패널요원에 대한 분산분석결과로 얻은 F값과 상관분석에 의한 상관계수는 Table 1에 있다.

F값은 패널요원이 다른 시료를 다르다고 평가하고 반복되는 똑같은 시료에 대해 일관성있게 평가하는 정도를 나타낸 수치⁽¹²⁾로서 F값이 높은 패널일수록 더 나은 패널요원이다⁽¹⁾. F값이 높은 패널요원은 비교적 넓은 범위의 평가점수를 사용하고 있다고 알려져 왔다. Cross 등⁽¹⁾은 각 항목별로 F값에 의한 순위를 정한 후, 각 항목별 F값의 순위합에 의해 패널요원의 수행능력을 평가하였다.

Table 1의 F값과 상관계수(r값)를 살펴보면 패널요원 1번, 2번, 11번은 모든 항목에 있어서 세 가지 다른 시료를 다르다고 평가했으며, 유의적인($p=0.05$) 상관계수를 보여주었다. 패널요원 7번과 9번은 모든 특성항목에 있어서 시료간의 차이를 감지하지 못했으며 상관계수 역시 유의적이 아니었다. 즉, 패널요원 1번, 2번, 7번, 9번, 11번의 경우 분산분석의 결과와 상관분석의 결과가 일치함을 보여주었다. 40명의 패널요원 중에서 패널요원 1번, 2번, 11번은 패널수행능력이 아주 뛰어난 패널요원들이었으며, 패널요원 7번과 9번은 패널성적이 아주 낫은 경우였다. 패널수행능성이 아주 양호하거나 아주 불량한 경우, 즉 양극단의 경우, 분산분석과 상관분석은 일치된 결과를 보여주었다.

Table 1. F ratios by analysis of variance and r values by correlation analysis of each panelist for each quality attribute

Panelist	Odor		Appearance		Flavor		Texture		Overall	Quality
	#	F ratio	r value	F ratio	r value	F ratio	r value	F ratio	r value	
1	54.6***	0.93***	15.9***	0.75**	18.2**	0.80**	19.0**	0.68*	12.2**	0.79**
2	25.8**	0.94***	9.0*	0.64*	10.1*	0.71**	7.0*	0.72**	13.0**	0.78**
3	0.6	0.39	12.5**	0.85***	10.5*	0.83***	3.4	0.77**	8.4*	0.85***
4	117.0***	0.97***	28.5***	0.95***	4.9	0.76**	1.4	0.44	6.0*	0.60*
5	6.5*	0.77**	2.7	0.28	0.3	0.27	45.0***	0.11	2.1	0.10
6	4.2	0.80**	37.0***	0.88***	3.6	0.59*	3.4	0.62*	13.0**	0.82***
7	0.1	0.20	0.6	-0.30	0.8	-0.17	1.4	0.14	0.8	-0.18
8	18.4**	0.82***	9.0*	0.79**	3.3	0.63*	2.8	0.64*	1.4	0.57
9	0.7	0.21	0.7	0.48	3.3	0.55	0.0	-0.10	0.4	0.37
10	9.8*	0.72**	5.9*	0.69*	10.8*	0.71**	10.8*	0.71*	3.0	0.69*
11	7.7*	0.77**	24.1**	0.86***	36.2***	0.91***	18.8**	0.89***	18.1*	0.90***
12	5.8*	0.75**	14.6**	0.80**	47.4***	0.55	6.5*	-0.07	11.5**	0.64*

*, **, ***Significant at p=0.05, 0.01, 0.001 respectively

패널요원 5번의 조직감, 패널요원 12번의 맛과 조직감 평가결과, F값은 p=0.05수준에서 유의적 이었지만 r값은 유의적(p=0.05)이 아니었다. 이 경우는 평가에 일관성을 지니고 다른 시료를 다르다고 평가 했지만, 전체패널의 평가점수와는 평가경향이 달름을 알 수 있다. 이런 패널요원은 평가기준, 혹은 평가절차를 잘못 이해하고 있는 경우이므로 재훈련이 요구된다. 패널요원의 수행능력을 F값의 순위에 의해 평가한다면, 패널요원 5번의 조직감, 패널요원 12번의 맛과 조직감의 평가결과와 같은 경우, 패널요원의 문제점을 파악하기 힘들다.

패널요원 3번의 조직감, 4번의 맛, 6번의 향, 맛, 조직감, 8번의 맛, 조직감, 10번의 종합적 품질평가결과, F값은 유의적(p=0.05)이지 못하였지만, r값은 유의적(p=0.05)이었다. 이들 패널요원들은 그 품질특성들을 평가하는데 일관성이 없었으며, 결과적으로 제시된 시료들은 통계적으로 차이가 없다고 했으나, 그들의 일반적인 평가경향은 전체패널의 평가경향과 상관이 있음(p=0.05)을 보여주고 있다. 이들 역시 각 품질특성에 대한 재훈련이 요구된다. 만약 상관분석에 의해 패널요원을 전체 패널 평가점수와 상관관계가 있는 요원과 상관이 없는 요원으로 나눈다면, 평가점수 경향과는 상관이 있으나 시료를 일관성있게 평가하지 못하는 패널요원을 가려낼 수 없는 단점이 있다.

주성분분석은 보통 패널의 점수에 의해 패널을 좀더 동질적인 그룹으로 나누는데 있다. 주성분 분석결과 제1주성분은 총변동의 65%를, 제2주성분은 총변동의 17%를 점유하였으며, 제1주성분과 제2주성분을 5개의 특성으로 아래와 같이 나타낼 수 있었다.

$$\text{제1주성분} = 0.35\text{향} + 0.40\text{외관} + 0.48\text{맛} + 0.49\text{조직감} + 0.49\text{전반적인 품질}$$

$$\text{제2주성분} = 0.83\text{향} + 0.19\text{외관} - 0.14\text{맛} - 0.48\text{조직감} - 0.15\text{전반적인 품질}$$

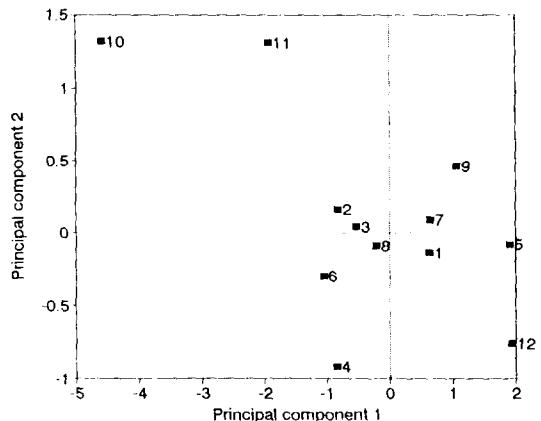


Fig. 1. Scores for 12 panelists on first (horizontal) and second (vertical) principal components from principal components analysis of sensory data

위의 식에서 제1주성분은 모두 양의 부호로 나타나 있고, 제2주성분에서 향과 외관은 양의 부호, 맛, 조직감, 전반적인 품질은 음의 부호로 나타나 있다. 여기에서 제2주성분은 입안에 직접 넣지 않고 판정 내릴 수 있는 특성과 입안에서 판정내릴 수 있는 특성을 양과 음의 부호로 나누었다고 볼 수 있지만 제1주성분과 제2주성분을 명쾌하게 해석하기가 곤란하며, 굳이 해석을 한다면 극히 주관적이며 임의적이기 쉽다.

제1주성분과 제2주성분에 의한 패널요원의 위치를 Fig. 1에 표시하였다. Table 1에서 패널요원중에서 패널요원 7번과 9번이 가장 수행능력이 떨어지게 평가되었지만 주성분분석 결과 이들은 다른 패널요원들과 비교적 비슷하게 나타났다(Fig. 1). 주성분분석 결과 전체적으로 패널요원 10번이 다른 패널들에 비해 많이 떨어져 있었고, 제1주성분 지표에서는 패널요원 5번, 11번, 12번이

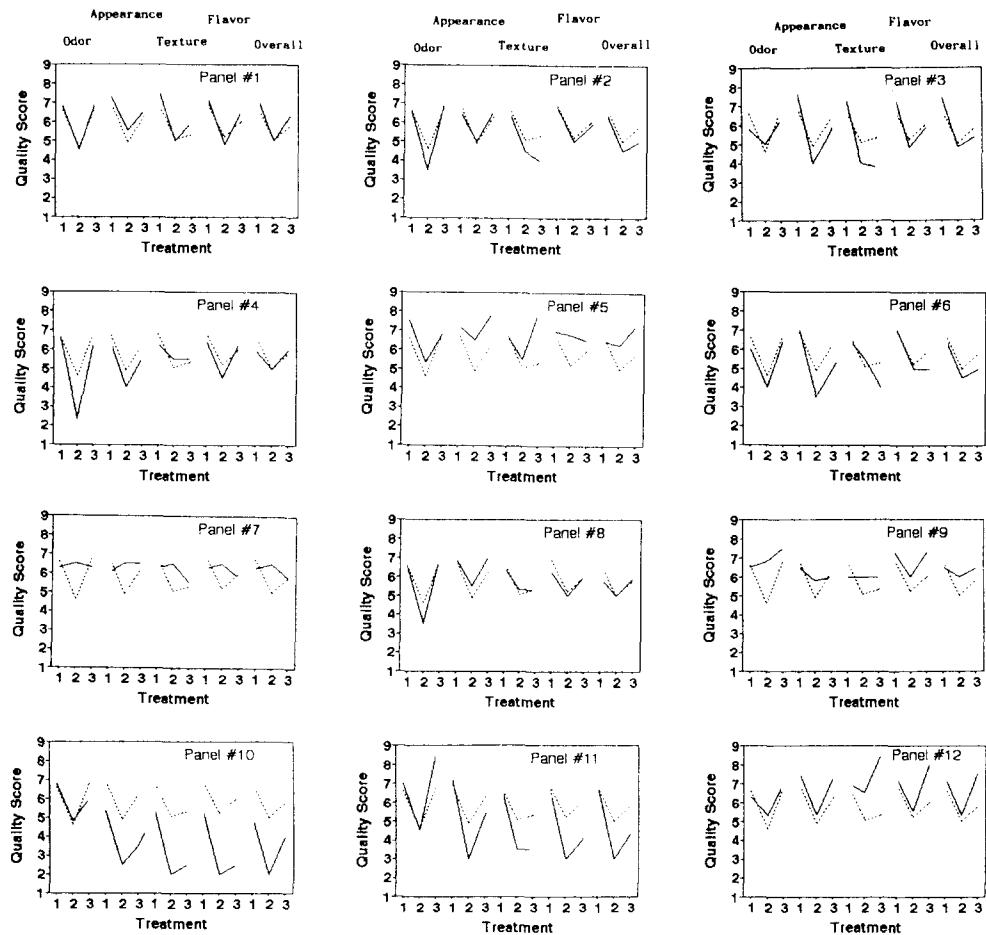


Fig. 2. Comparison of mean scores of each panelist and those of whole panel for each attributes depending on the treatments

(---; mean scores of whole panel, —; mean scores of each panelist)

원점에서 가장 많이 떨어져 있었으며, 제2주성분 지표에서는 패널요원 11번, 4번, 12번 이 다른 패널요원들 보다는 많이 떨어져 있었다.

각 관능특성항목에 따른 패널의 평균점수와 비교한 각 패널요원의 품질평가점수는 Fig. 2에 있다. 패널요원 10번은 Table 1에서 각 특성항목별 F 값과 r 값을 살펴보면 5% 수준에서 제품들간의 향, 외관, 맛 및 조직감에서의 차를 감지 하였고 Fig. 2에서는 평가점수경향은 평균점수와 같았으나 사용한 점수범위가 다른 어느 패널요원에 비해 매번 일관성있게 아주 낮았다.

관능검사 결과에 대한 패널과 검사시료의 효과를 보기위해 이원 분산분석을 하는데, 이때 패널과 검사물의 상호작용(interaction)이 유의적인 경우 패널들이 검사물을 평가하는데 차이가 있음을 나타낸다. 상호작용에는 2가지 종류가 있는데 첫번째는 전체 패널과 개개인 패널이 제품들간의 특성강도 평가에는 같은 경향을 보이나

평가점수 범위가 달라서 제품과 특성강도 그래프에서 간격이 벌어진, 기울기가 같지않은 직선들로 나타나는 경우와, 두번째는 전체 패널과 개개인 패널 사이의 평가경향이 일치하지 않아 X자 모양으로 평가반응이 교차하는 경우가 있다^[14]. 전자는 단지 패널의 특성 강도에 대한 민감성과 평가점수 범위의 차이를 반영하여, 전체적인 제품의 차이에는 영향을 주지 않는 반면, 후자의 경우는 제품간 차이의 유의성을 상실하게 될 수도 있다. 주성분분석 결과 전체적인 제품의 차이에 영향을 주는 평가경향이 다른 패널요원 7번과 9번 보다는 전체적인 제품의 차이에 영향을 주지 않는 즉, 평가경향은 같지만 평가점수 범위가 다른 패널요원 10번이 outlier로 나타났다. 주성분분석을 이용하여 패널수행능력을 평가할 경우 제품의 차이를 감지 못하는 패널요원보다는 단지 사용된 점수범위가 다른 패널요원을 outlier로 판정내릴 수 있다는 것을 보여주었다. Powers^[9]는 다변량분석방

법만으로 패널을 선정하는 것은 충분하지 않으며 분산분석과 함께 사용되어야 한다고 하였으며, 본 연구에서도 주성분분석 하나만으로 패널수행능력을 평가하는 것은 평가경향이 전혀 다른 패널을 선별해 낼 수 없는 단점이 있는 것으로 나타났다. 그러므로 주성분분석은 자료분석의 마지막 단계이기보다는 차후 분석에 필요한 새로운 지표를 얻기위한 중간단계로 사용됨이 바람직하다.

결론적으로 각 패널요원의 수행능력을 평가하기 위해 서는 분산분석이나 상관분석 중의 한 방법만을 사용하기보다는 두 가지 방법을 모두 사용한다면 좀 더 각 패널요원들에 대한 장단점을 파악하게 되고 재훈련 과정 동안에 각 패널요원의 약한 부분을 정확하게 보완해 줄 수 있을 것이다. 또한 주성분분석만을 패널평가 방법으로 사용할 경우에는 제품의 차이를 감지 못하는 패널보다는 평가점수 범위가 전체패널과 다른 패널을 outlier로 선정할 위험이 있어 패널평가방법으로는 적합하지 않은 것으로 나타났다.

문 현

1. Cross, H.R., Moen, R. and Stanfield, M.S.: Training and testing of judges for sensory analysis of meat quality. *Food Technol.* 32(7), 48(1978)
2. Gacula, M.C. and Singhi, J.: Statistical methods in food and consumer research. Chapter 5. Incomplete Block Designs. Academic Press, San Diego, CA. p.141(1984)
3. Malek, D.M., Munroe, J.H., Schmitt, D.J. and Korth, B.: Statistical evaluation of sensory judges. *Am. Soc. Brew. Chem. J.* 49(1), 23(1986)
4. McDaniel, M., Henderson, L.A., Watson, JR. B.T. and

- Heatherbell, D.: Sesnory panel training and screening for descriptive analysis of the aroma of Pinot Noir Wine fermented by serveral strains of malolactic bacteria. *J. Sensory Studies.* 2(3), 149(1987)
5. Ennis, D.M.: Multivariate sensory analysis. *Food Technol.* 42(11), 118(1988)
 6. Ennis, D.M., Boelens, H., Haring, H. and Bowman, P.: Multivariate analysis in sensory evaluation. *Food Technol.* 36(11), 83(1982)
 7. 허명희: 주성분분석과 인자분석. In: 농림수산 연구조사설계의 최적화와 분석법. 제1회 생물통계연구회 주최워크숍. 서울대학교 농과대학. 수원(1993)
 8. Wu, L.S., Bargmann, R.E. and Powers, J.L.: Factor analysis applied to wine descriptors. *J. Food Sci.* 42, 944(1977)
 9. Powers, J.J.: Uses of multivariate methods in screening and training sensory panelists. *Food Technol.* 42 (11), 123(1988)
 10. Malundo, T.M.M. and Resurrecion, A.V.A.: A comparison of performance of panel selected using analysis of variance and cluster analysis. *J. Sensory Studies.* 7(1), 63(1992)
 11. Gatchalian, M.M., Leon, S.Y.D. and Yano, T.: Control chart technique, a feasible approach to measurement of panelist performance in product profile development. *J. Sensory Studies.* 6(4), 239(1991)
 12. American Society of Testing and Materials. Guidelines for the selection and training of sensory panel members. *ASTM Special Technical Publication* 758(1981)
 13. SAS Institute : SAS/Stat User's Guide, release 6.03 ed., Cary, NC. SAS Institute Inc.(1988)
 14. Stone, H. and Sidel, J.L.: Sensory evaluation practices. Academic press, San Diego, CA. (1985)

(1993년 11월 16일 접수)