# SURVIVAL ANALYSIS TO ESTIMATE SENSORY SHELF LIFE USING ACCEPTABILITY SCORES

ANA GIMÉNEZ, GASTÓN ARES[1] and ADRIANA GÁMBARO

*Sección Evaluación Sensorial*
*Departamento de Ciencia y Tecnología de Alimentos*
*Facultad de Química*
*Universidad de la República*
*Gral. Flores 2124. C.P. 11800, Montevideo, Uruguay*

Accepted for Publication September 20, 2007

## ABSTRACT

*In the present study, survival analysis was applied on consumers' acceptability scores, focusing the shelf-life risk on the consumers disliking the product. Shelf life was estimated as the time necessary to reach a fixed percentage of consumers disliking the product, that is, the percentage of consumers scoring the product's overall acceptability below 6 in a 9-point hedonic scale. In the present work, this methodology was applied to estimate the sensory shelf life of whole pan bread and alfajores. Shelf lives estimated considering 50% of the consumers disliking the product were shorter than those estimated considering consumers' rejection to consume. These results suggest that a proportion of consumers might dislike the sample but still answer "yes" when asked if they would consume the product at their homes. This could be attributed to the fact that when consumers are asked whether they would consume the product, they might think about consuming it after being stored at their homes. In this situation, consumers might be more tolerant toward sensory defects because they do not want to discard the product.*

## PRACTICAL APPLICATIONS

Sensory shelf life can be estimated using survival analysis considering the percentage of consumers disliking the product. This methodology seems to be a conservative criterion in order to assure the product's quality throughout its storage.

[1] Corresponding author. TEL: 5982-9245735; FAX: 5982-9241906; EMAIL: gares@fq.edu.uy

 

## INTRODUCTION

The shelf life of many products is determined by their sensory shelf life. Methodologies to estimate sensory shelf life are necessary in food product and process development in order to study the influence of changes in formulation and manufacturing processes. Furthermore, accurate shelf-life labeling is necessary in order to extend commercialization times to its maximum but assuring the product's freshness. Finding unacceptable products within their shelf life could diminish consumers' confidence in a brand and in the store that sells it (Harcar and Karakaya 2005).

Traditionally, the sensory shelf life of food products was determined as the time required for a certain sensory attribute to reach a predetermined intensity, usually called failure criteria (Hough *et al.* 2002). Traditionally, failure criteria have been arbitrarily selected without any studies to support their validity or to correlate these sensory limits with consumers' perception.

However, food products do not have sensory shelf lives of their own; rather, they depend on the interaction of the food with the consumer (Hough *et al.* 2003). Consumers are the ones who decide if they would accept a food product after a certain storage time. For this reason, consumers are the most appropriate tool for determining a food product's sensory shelf life (Hough *et al.* 2003).

Different methodologies could be used to determine the shelf life of a food product using consumer data. Sensory shelf life could be determined as the time required for the overall acceptability scores of the product to fall below a certain predetermined value – or failure criteria – for example, a value of 6.0 in a 9-point structured hedonic scale (Muñoz *et al.* 1992; Giménez *et al.* 2007). In this methodology, failure criteria are selected based on consumers' perception of the product rather than on trained assessors' data.

In order to estimate sensory shelf life based on consumers' rejection of a food product, survival analysis could have been applied (Hough *et al.* 2003). This methodology focuses the shelf-life risk on the consumers' rejection of the product, estimating shelf life as the time necessary to reach a fixed percentage of consumer rejection, traditionally 25 and 50% (Hough *et al.* 2003; Gámbaro *et al.* 2005, 2006a; Ares *et al.* 2006; Giménez *et al.* 2007). These percentages mean that if a consumer tries the product at the end of its shelf life, there is a 25 or 50% probability that he will reject it. However, this could not be conservative enough as to assure the product's freshness and quality near the end of its shelf life. During their decision-making process, consumers rely on different attributes before deciding whether to buy or consume a certain food product (Gardial *et al.* 1994; Ragaert *et al.* 2004). Besides, evaluative criteria may change depending on the stage of the decision-making process, whether it is right after purchase or at home after the product has been stored in their refrigerators. For this reason, consumers might be more tolerant to products' defects if they are thinking of consuming a product that was stored at their homes than if they are thinking of consuming a product immediately after purchase. Thus, estimated shelf life may vary if consumers are asked whether they are imagining consuming the product right after purchase or after being stored at their homes.

Food producers might be interested in selling products within their shelf lives, and that consumers actually like them. However, consumers might accept to consume a product for its consumption but not actually enjoy it. Therefore, a methodology to assure that consumers actually like the product at the end of its shelf life is necessary.

In the commonly used 9-point hedonic scale, a score of 5 indicates "neither like nor dislike." Thus, 6 is the first score indicating that the consumer likes the product. For this reason, 6–9 scores could be regarded as liking scores. Therefore, survival analysis could be used focusing the shelf-life risk on the consumers disliking the product, estimating shelf life as the time necessary to reach a fixed percentage of consumers disliking the product, for example, 50%.

The aims of the present work were to (1) evaluate the application of survival analysis to estimate sensory shelf life using acceptability scores; and (2) compare results with those obtained using acceptability limit methodology and survival analysis with rejection to consume data.

In the present article, data from Gámbaro *et al.* (2005, 2006b) are reanalyzed.

## MATERIALS AND METHODS

### Samples

Two studies were carried out, one for evaluating the sensory shelf life of whole pan bread, and one for alfajores.

A commercial whole pan bread was used. Breads packaged in polyethylene bags were bought from a local distributor, all from the same batch. The breads were stored in a temperature-controlled storage room at 20C for 1, 4, 7, 10, 13, 15 and 17 days. After reaching the desired storage times, the breads were frozen at −20C and stored at −18C, providing samples with different storage times (Gacula and Kubala 1975). The samples were defrosted at 20C for 6 h for its evaluation.

Alfajor is a chocolate-coated individually wrapped cake that is popular in some Latin American countries such as Argentina and Uruguay. It is defined by Uruguayan legislation as a product obtained by joining two or more layers of cookies or cake with filling, usually dulce de leche, covered or enrobed with various confectionary coatings, chocolate being the most common (M.S.P. 1994). Alfajores used in the experiment were provided by a local manufacturer. Storage times at 35C were 0, 10, 16, 25, 30, 35, 39 and 45 days. Samples were manufactured for each of the sampling times and were stored at 35C. As the samples belonged to different batches, the trained sensory panel performed triangular tests to ensure that there was no significant difference between a sample of the previous batch (stored at 5C) and that of the following batch.

Storage times were selected based on previous studies (Gámbaro *et al.* 2005; Giménez *et al.* 2007).

## Consumer Study

Two consumer studies were performed, one for evaluating whole pan bread and the other for evaluating alfajores.

In each study, 50 people who consumed each product, aged between 18 and 50, were recruited from the city of Montevideo, Uruguay. Consumers received the seven samples corresponding to the seven storage times of one product, following a balanced complete block design. The consumers received one-half bread loaf or one-fourth of alfajor in closed plastic containers coded with three-digit random numbers.

For each sample, they had to score its overall acceptability using a 9-box scale labeled on the left with "dislike very much," in the middle with "neither like nor dislike," and on the right with "like very much." They also answered the question "would you normally consume this product?" with a "yes" or a "no."

The testing was carried out in a sensory laboratory that was designed in accordance with the ISO (1988).

## Data Analysis

**Regression Analysis.** A linear regression was carried out considering consumers' overall acceptability as a dependent variable and storage time as an explanatory variable. Using this regression, sensory shelf life could also be determined as the time required for the overall acceptability scores of the product to reach a certain predetermined value or failure point. Muñoz *et al.* (1992) considered an acceptability score of 6.0 in a 9-point hedonic scale as the commercial or quality limit for food products. Therefore, this criterion was used for estimating the sensory shelf life of the evaluated products.

**Survival Analysis Using Consumers' Rejection.** Survival analysis methodology was used to estimate the shelf life of whole pan bread and alfajores using results obtained from consumers when asked if they would normally consume the samples with different storage times.

Defining a random variable $T$ as the storage time at which a consumer rejects the sample, the rejection function $F(t)$ can be defined as the probability of a consumer rejecting a product before time $t$, that is $F(t) = P(T \leq t)$ (Hough *et al.* 2003).

A parametric model can be used to estimate the rejection function and other quantities of interest (Hough *et al.* 2003). For example, if a Weibull distribution is chosen for $T$ (Klein and Moeschberger 1997; Lindsay 1998), the rejection function is given by

$$F(t) = 1 - S_{\text{sev}}\left(\frac{\ln(t) - \mu}{\sigma}\right)$$

where $S_{\text{sev}}(\cdot)$ is the survival function of the smallest extreme value distribution $S_{\text{sev}}(w) = \exp(-e^{w})$, and $\mu$ and $\sigma$ are the model's parameters.

To estimate shelf life, the probability of a consumer rejecting a product (that is, $F[t]$) must be chosen. The most commonly used consumer rejection percentages are 25 and 50% (Hough *et al.* 2003; Gámbaro *et al.* 2006a), and therefore these percentages were considered in the present work. Ninety-five percent confidence intervals for sensory shelf lives were also calculated and were used to compare the different estimated shelf lives. If confidence intervals do not overlap, two sensory shelf life values are significantly different, and if they do, they are not significantly different.

Calculations were performed using procedures from S-PLUS statistical software (Insightful Corp., Seattle, WA). A 5% significance level was considered.

**Survival Analysis Using Consumers' Acceptability Scores.** In order to estimate sensory shelf life using acceptability scores, a data transformation was performed. If the acceptability score of a consumer for a sample was 1–5, then the rating was transformed to the word "no," indicating that the consumer disliked the product. In contrast, if a consumer's score for a sample was 6–9, it was replaced by the word "yes," indicating that the consumer liked the product.

Therefore, if survival analysis is performed on transformed data, the probability of a consumer disliking the product as a function of time could be calculated. A random variable $S$ could be defined as the storage time at which the consumer stops liking the sample. Thus, function $G(t)$ can be defined as the probability of a consumer disliking a product before time $t$, that is $G(t) = P(S \leq t)$. Data were analyzed as with rejection to consume data.

To estimate shelf life, the probability of a consumer disliking a product (that is, $G[t]$) must be chosen. In this case, 25 and 50% were considered. These percentages were selected considering that a food producer might want that most of the consumers like his product. Thus, the selected percentages mean that if a consumer tries the product at the end of its shelf life, there is a 25 or 50% probability that he will dislike it. Only few consumers usually taste the product at the end of its shelf life. Thus, considering this and the fact that of the few that do, 75 or 50% will still find the product acceptable, the value of $G(t) = 25$ or 50% is considered reasonable from a practical point of view as to assure the product's quality. Ninety-five percent confidence intervals for sensory shelf lives were also calculated and were used to compare the different estimated shelf lives.

Calculations were performed using procedures from S-PLUS statistical software (Insightful Corp.). A 5% significance level was considered.

## RESULTS

### Shelf-Life Estimation Using an Acceptability Limit

As shown in Fig. 1, initial acceptability scores for whole pan bread and alfajores were higher than 7.0. Muñoz *et al.* (1992) considered an acceptability score of 6.0 in a 9-point hedonic scale as commercial or quality limit. Therefore, shelf life could be estimated as the storage time when acceptability reaches a value of 6. A linear regression was performed on acceptability scores versus storage time. This regression gave a good fit for both whole pan bread and alfajores ($R^2 = 0.931$ and 0.908, respectively). Using this curve, sensory shelf lives could be estimated in $9 \pm 1$ days for whole pan bread and $15 \pm 2$ days for alfajores.

### Survival Analysis Using Rejection to Consume Data

Because there are no statistical tests to compare the goodness of fit of the different parametric models used for interval-censored data, visual assessment of how the parametric models adjust to the nonparametric estimation was used to choose the most adequate model. The following distributions were considered: smallest extreme value, normal, logistic, Weibull, log normal and log logistic (Hough *et al.* 2003). As the Weibull distribution adjusted best for both products, it was chosen to model rejection times for the present data.

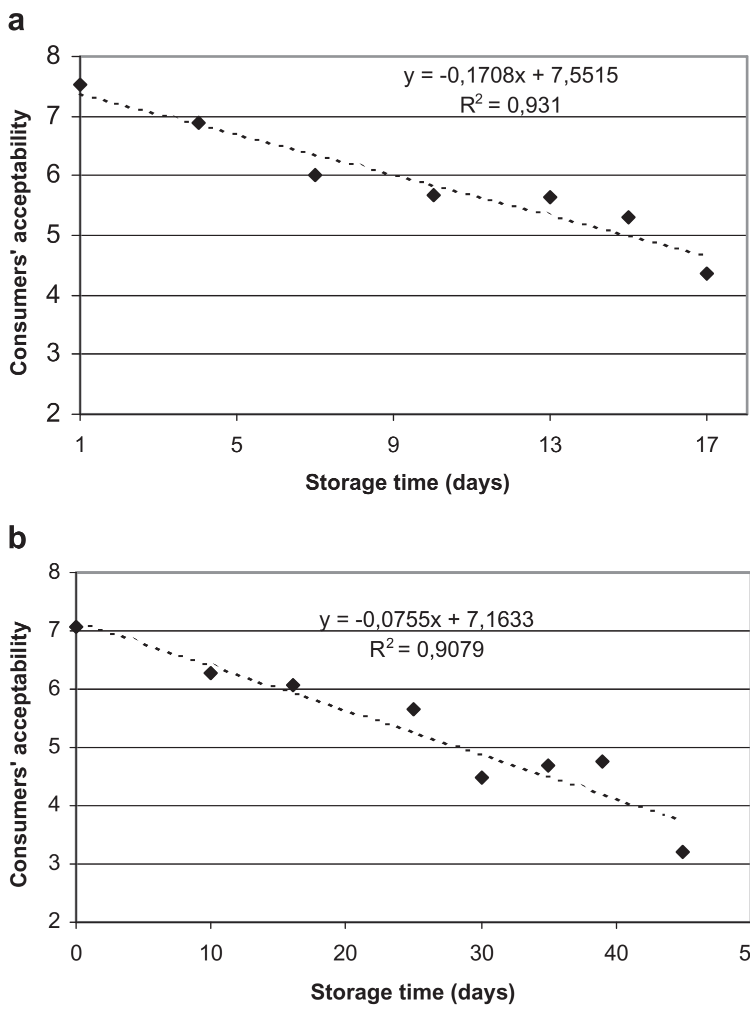


FIG. 1. AVERAGE CONSUMERS' ACCEPTABILITY SCORES AS A FUNCTION OF STORAGE TIME FOR WHOLE PAN BREAD AND ALFAJORES

The maximum likelihood estimates of the parameters of the Weibull distribution for each product are shown in Table 1. These parameters were used to graph the percentage of consumers' rejection to consume versus storage time for each product, as shown in Fig. 2.

Shelf lives were calculated by estimating the time necessary to achieve a certain consumer rejection percentage, in this case 25 and 50%. Shelf-life estimations are shown in Table 2.

TABLE 1.
VALUES OF WEIBULL DISTRIBUTION PARAMETERS μ AND σ AND THEIR CONFIDENCE INTERVALS FOR THE FAILURE FUNCTION, CONSIDERING REJECTION-TO-CONSUME AND ACCEPTABILITY DATA, FOR WHOLE PAN BREAD AND ALFAJORES

| Product | Rejection to consume | | Acceptability scores | |
|---|---|---|---|---|
| | μ ± standard error | μ ± standard error | σ ± standard error | σ ± standard error |
| Whole pan bread | 2.900 ± 0.042 | 0.396 ± 0.034 | 2.653 ± 0.027 | 0.445 ± 0.030 |
| Alfajores | 3.752 ± 0.038 | 0.212 ± 0.040 | 3.462 ± 0.034 | 0.382 ± 0.024 |

TABLE 2.
ESTIMATED SHELF LIVES (DAYS) (VALUE ± STANDARD ERROR) CONSIDERING REJECTION-TO-CONSUME DATA AND PERCENTAGE OF CONSUMERS DISLIKING THE PRODUCT, FOR WHOLE PAN BREAD AND ALFAJORES

| Product | Rejection to consume | | Consumers disliking the product | |
|---|---|---|---|---|
| | 25% | 50% | 25% | 50% |
| Whole pan bread | 11 ± 1 | 16 ± 2 | 8 ± 1 | 12 ± 1 |
| Alfajores | 33 ± 2 | 39 ± 3 | 20 ± 2 | 28 ± 2 |

## Survival Analysis Using Acceptability Scores

In order to select between the different parametric models to describe the data, the visual procedure previously described was performed. The Weibull distribution adjusted best for both products, so it was chosen to model dislike times for the present data.

The maximum likelihood estimates of the parameters of the Weibull distribution for each product are shown in Table 1. These parameters were used to graph the percentage of consumers disliking the product versus storage time for each product, as shown in Fig. 2.

Shelf lives were calculated by estimating the time necessary to achieve 25 and 50% of the consumers disliking the product, as shown in Table 2.

# DISCUSSION

Although the samples of both products showed low average acceptability scores at the end of the evaluated storage time, a high proportion of the consumers would still consume them. Thus, although the consumers might not like the product, they would consume it anyway. For example, although

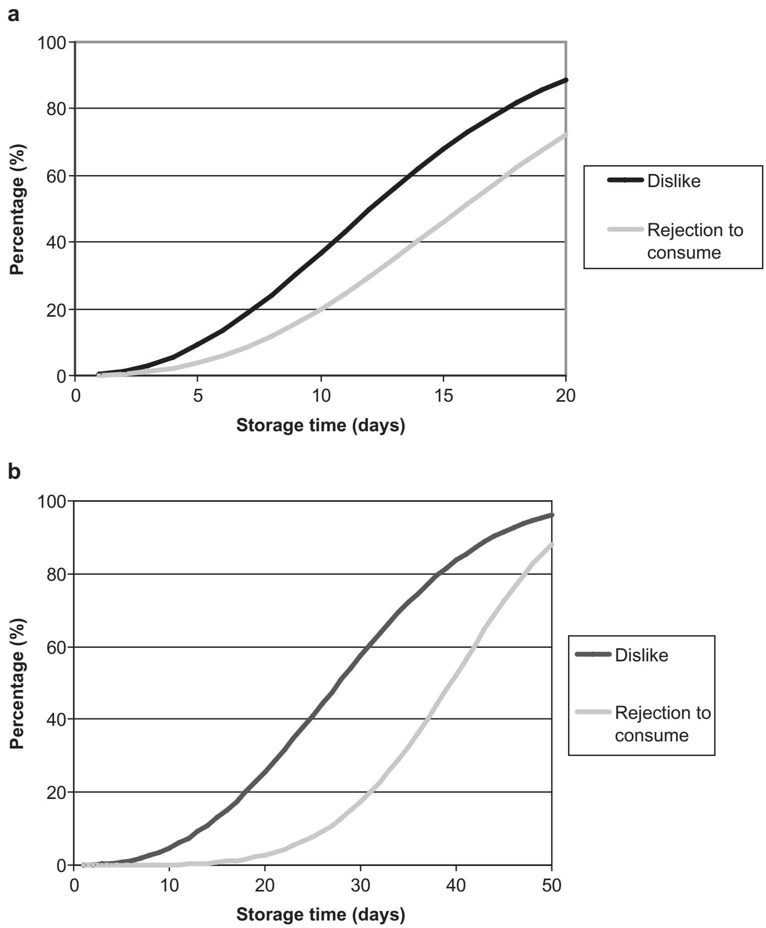


FIG. 2. PERCENTAGE OF CONSUMERS' REJECTION TO CONSUME AND PERCENTAGE OF CONSUMERS DISLIKING THE SAMPLE AS A FUNCTION OF STORAGE TIME FOR WHOLE PAN BREAD AND ALFAJORES

alfajores stored for 30 days showed an average accepatiblity score of 4.6, only 20% of the consumers rejected to consume them. This trend was also reflected in the percentage of consumers disliking the samples and in the percentage of consumers' rejection to consume. As shown in Fig. 2, throughout the storage, the percentage of consumers disliking the sample was higher than the percentage of consumers rejecting them, for both products. These results suggest that a proportion of the consumers might dislike the sample but still answer "yes" when asked if they would consume the product at their homes, indicating the influence of context and proper question formulation on shelf-life estimations.

The difference between rejection to consume and consumers disliking the product could be attributed to the fact that when consumers think about consuming the product after being stored at their homes, they are more tolerant to defects because they have already bought the product and they do not want to discard it. Although they do not actually like the product, they consume it anyway. However, if they buy the product and immediately consume it, they might dislike the product and might be dissapointed. Thus, considering the percentage of consumers disliking the sample seems to be a more conservative criterion in order to assure the product's quality at the end of its shelf life.

As shown in Table 2, shelf-life estimations considering consumers' rejection to consume were larger than those estimated considering the percentage of consumers disliking the product. Shelf lives estimated considering 50% of the consumers disliking the product were close to those estimated considering 25% consumers' rejection, suggesting that both criteria might provide similar results.

On the other hand, sensory shelf lives estimated considering 50% rejection to consume were significantly longer than those estimating considering 50% of the consumers disliking the product. Thus, estimating sensory shelf life as the time necessary to reach 50% consumers' rejection to consume may imply that most of the consumers who try the product might dislike it (73 and 82% for whole pan bread and alfajores, respectively). Consumers' overall acceptability scores corresponding to 50% consumers' rejection were 4.8 and 4.2 for whole pan bread and alfajores, respectively. These results suggest that if shelf lives are estimated considering 50% consumers' rejection to consume, a large proportion of the consumers who try the product near the end of its shelf life might find it unacceptable, could be dissapointed and complain, or even not buy the product again. This could diminish consumer confidence in the brand or in the store that sells it. Therefore, the shelf life estimated considering 50% consumers' rejection seems to be not conservative enough.

Sensory shelf lives estimated considering 25% rejection to consume were shorter than that estimated considering an acceptability limit of 6.0. The difference was larger in the case of alfajores, suggesting that the relationship between acceptability scores and consumers' rejection to consume may be affected by the product considered.

Sensory shelf lives estimated considering an acceptability limit of 6.0 were shorter than those estimated considering 50% of the consumers disliking the product. Whereas in the case of alfajores this estimation was also shorter than that for 25% of the consumers disliking the product, in the case of whole pan bread, both estimations were similar. These differences could be attributed to the fact that consumers' reactions are heterogeneous, and averaging their responses might cause an important loss of data, ignoring groups of consumers with different sensitivity toward products' defects. The acceptability limit methodology might be too strict a criterion in some cases. One of the main advantages of survival analysis is the fact that individual responses are considered for shelf-life estimations, in agreement with the fact that sensory shelf life depends on the interaction of the product and the consumer.

## CONCLUSIONS

The shelf life of whole pan bread and alfajores was estimated as the time necessary for 25 and 50% of the consumers to dislike the product, that is, to score its acceptability below 6 in a 9-point hedonic scale. Shelf-life estimations using this methodology were shorter than those estimated considering 50% consumers' rejection to consume. Thus, considering 50% of the consumers disliking the sample seems to be a more conservative criterion in order to assure the product's quality at the end of its shelf life. This could be attributed to the fact that when consumers are asked whether they would consume the product, they might think about consuming the product after being stored at their homes. In this context, they are more tolerant to defects because they have already bought the product and they do not want to throw it away, rather thinking of consuming the product right after purchase. Sensory shelf lives estimated considering 25% consumers' rejection to consume were similar to those estimated considering 50% of the consumers disliking the product, suggesting that the former would be a more valid criterion than 50% consumers' rejection to consume.

Further work is needed in order to evaluate the influence of context on sensory shelf life of other type of products, and to address some of the limitations of the present study, such as the small consumer sample.

## REFERENCES

ARES, G., PARENTELLI, C., GÁMBARO, A., LAREO, C. and LEMA, P. 2006. Sensory shelf life of shiitake mushrooms stored under passive modified atmosphere. Postharvest Biol. Technol. *41*, 191–197.

GACULA, M.C., JR. and KUBALA, J.J. 1975. Statistical models for shelf-life failures. J. Food Sci. *40*(2), 404–409.

GÁMBARO, A., GIMÉNEZ, A., VARELA, P., GARITTA, L. and HOUGH, G. 2005. Sensory shelf-life estimation of alfajor by survival analysis. J. Sensory Studies *19*, 500–509.

GÁMBARO, A., ARES, G. and GIMÉNEZ, A. 2006a. Shelf life estimation of apple baby food. J. Sensory Studies *21*(1), 101–111.

GÁMBARO, A., ARES, G., GIMÉNEZ, A. and GILARDI, V. 2006b. Influence of enzymes on the texture of brown pan bread. J. Texture Studies *37*, 300–314.

GARDIAL, S.F., CLEMONS, D.S., WOODRUFF, R.B., SCHUMANN, D.W. and BURNS, M.J. 1994. Comparing consumer's recall of prepurchase and postpurchase product evaluation experiences. J. Consum. Res. *20*, 548–560.

GIMÉNEZ, A., VARELA, P., SALVADOR, A., ARES, G., FISZMAN, S. and GARITTA, L. 2007. Shelf life estimation of brown bread: A consumer approach. Food Qual. Prefer. *18*, 196–204.

HARCAR, T. and KARAKAYA, F. 2005. A cross-cultural exploration of attitudes toward product expiration dates. Psychol. Market. *22*(4), 353–371.

HOUGH, G., SÁNCHEZ, R.H., GARBARINI DE PABLO, G., SÁNCHEZ, R.G., CALDERÓN VILLAPLANA, S., GIMÉNEZ, A.M. and GÁMBARO, A. 2002. Consumer acceptability versus trained sensory panel scores of powdered milk shelf-life defects. J. Dairy Sci. *85*(9), 1–6.

HOUGH, G., LANGOHR, K., GÓMEZ, G. and CURIA, A. 2003. Survival analysis applied to sensory shelf life of foods. J. Food Sci. *68*, 359–362.

ISO. 1988. *Sensory Analysis. General Guidance for the Design of Test Rooms*, ISO 8589:1988, ISO, Geneva, Switzerland.

KLEIN, J.P. and MOESCHBERGER, M.L. 1997. *Survival Analysis: A Self Learning Text*, Springer-Verlag, New York, NY.

LINDSAY, J.K. 1998. A study of interval censoring in parametric regression models. Lifetime Data Anal. *4*, 329–354.

M.S.P. (MINISTERIO DE SALUD PÚBLICA). 1994. *Reglamento Bromatológico Nacional, Decreto 315/994*, Ministerio de Salud Pública, Montevideo, Uruguay.

MUÑOZ, A.M., CIVILLE, V.G. and CARR, B.T. 1992. *Sensory Evaluation in Quality Control*, Van Nostrand Reinhold, New York, NY.

RAGAERT, P., VERBEKE, W., DEVLIEGHERE, F. and DEBEVERE, J. 2004. Consumer perception and choice of minimally processed vegetables and packaged fruits. Food Qual. Prefer. *15*, 259–270.