Survival Analysis Applied to Sensory Shelf Life of Foods

G. HOUGH, K. LANGOHR, G. GÓMEZ, AND A. CURIA

ABSTRACT: Survival analysis concepts to be used in sensory shelf life studies were introduced, together with the equations necessary for calculations. The survival function was defined as the probability of consumers accepting a product beyond a certain storage time. Censoring phenomena, a key concept in survival analysis, was defined and has been shown to occur in sensory shelf-life data. Concepts and calculations were applied to a data set obtained from 50 consumers who each tasted seven yogurt samples with different storage times, answering "yes" or "no" to whether they would consume the samples. From this censored data set, nonparametric and parametric models were obtained that allowed shelf-life estimations.

Keywords: shelf-life, sensory, survival analysis, statistics

Introduction

Sensory evaluation is the key factor for determining the shelf life of many food products. Microbiologically stable foods, such as biscuits or mayonnaise, will have their shelf life defined by the changes in their sensory properties. Many fresh foods, such as yogurt or pasta, after relatively prolonged storage may be microbiologically safe to eat but may be rejected due to changes in their sensory properties.

Survival analysis (Meeker and Escobar 1998; Klein and Moeschberger 1997; Kleinkamnbaun 1996; Gómez 2002) is a branch of statistics used extensively in clinical studies, epidemiology, biology, sociology, and reliability studies. Gacula and Singh (1984) introduced the Weibull model, derived from survival analysis, in shelf-life studies of food; later, the model was applied in some studies (Hough and others 1999; Cardelli and Labuza 2001; Duyvesteyn and others 2001). Censoring phenomena, a key concept in survival analysis and defined further ahead, has not yet been considered.

Traditionally the shelf life of foods has centered on the product. For example, O’Conner and others (1994) report that the shelf life of minimally processed kiwifruit stored at 4 °C was 2 d, based on flavor and texture changes measured with a trained sensory panel. Thus, in this study the hazard or the risk was focused on the fruit. But some kiwifruit consumers would probably have accepted this kiwifruit stored for 2 d, and it is also probable that another group of consumers would have rejected the fruit stored for only 1 d. Thus, from a sensory point of view, food products do not have shelf lives of their own, rather they will depend on the interaction of the food with the consumer. In survival analysis, the survival function S(t) is defined as the probability of an individual surviving beyond time t (Klein and Moeschberger 1997; Gómez 2002). Referring this definition to the sensory shelf life of the kiwifruit, the “individual” would not be the fruit itself, but rather the consumer, that is the survival function would be defined as the probability of a consumer accepting a product stored beyond time t. The hazard would not be focused on the product deteriorating, rather on the consumer rejecting the product.

Vaisey-Gensler and others (1994) used logistic regression analysis to relate the average proportion of acceptance of canola oils to their storage time. An incomplete block design was used whereby not all consumers tasted all storage times. This introduced a certain amount of confounding to the experiment; for example, the average proportion of acceptance at d 10 could be different to that of d 12 due to different storage times or due to different consumers. This type of incomplete block designs could have been analyzed more efficiently considering survival analysis concepts. Cardelli and Labuza (2001) used consumer information to define the sensory shelf life of coffee stored between 0 and 23.3 wk. At wk 0, 3 consumers were used. As the storage time increased, the number of consumers used was also increased. Thus, for example, at 20.1 wk of storage, 5 new consumers were recruited, of which 4 rejected the coffee and 1 accepted it. The rejection storage time for these 4 consumers was not T = 20.1 wk, rather T < 20.1 wk. That is, all we know is that they rejected the coffee with 20.1 wk storage, but we do not know if they would have rejected or accepted coffee stored for, for example, 15 wk. This uncertainty in the time of interest, as explained in the next section, is defined as censored data in survival analysis.

In the present work, survival analysis concepts to be used in sensory shelf-life studies are introduced, together with the necessary equations to carry out the calculations. A data set for stirred yogurt is presented as an example.

Survival Analysis Concepts

Definitions. Methods of survival analysis have been developed to evaluate times until an event of interest, often called survival times, taking into account the presence of censored data (Gómez 2002). These occur whenever a time of interest—for example, the time until death in epidemiological studies or the time until failure of machines in reliability studies—cannot be observed exactly. If we know this time to be superior to an observed time, we speak of right-censored data; if it is known to be less than the observation time, the time is called left-censored. Interval-censoring is given when the time of interest falls into an observed interval.

In food shelf-life studies, samples with different storage times are presented to consumers. Assume that we define a random variable T as the storage time on which the consumer rejects the sample. Thus, the survival function S(t) can be defined as the probability of a consumer accepting a product beyond time t, that is S(t) = P(T > t).

To illustrate the censored nature of the data, suppose that the storage times are: 0, 5, 10, and 15 d. Within this illustration, and
because of the discrete nature of the storage times, \( T \) will never be observed exactly, and instead we will observe whether \( T \leq 5 \), \( a \leq T \leq b \), or \( T \geq 15 \). This incompleteness of the data can be related to usual censoring definitions in survival analysis:

Left censoring: If a consumer rejects the sample with 5 d storage, time to rejection is \( \leq 5 \) d. This is a consumer who is very sensitive to storage changes, and rejects the product somewhere between 0 and 5 d.

Interval censoring: If, for example, a consumer accepts samples stored at times 0 and 5 d, but rejects the sample stored at 10 d, the time to rejection is \( 5 < T \leq 10 \) d. Our resources, or practical considerations, did not allow presenting the consumer with samples stored at 6, 7, 8, and 9 d; thus, what we know is that the consumer would reject the product stored between 5 and 10 d.

Right censoring: If a consumer accepts all samples, we would say that time to rejection is \( > 15 \) d. That is, we suppose that, if a sample is stored a sufficiently long time, the consumer will eventually reject it.

Nonparametric estimation of the survival function. The likelihood function, which is used to estimate the survival function, is the joint probability of the given observations of the \( n \) consumers (Klein and Moeschberger 1997):

\[
L = \prod_{i=1}^{n} S(t_i)^{I_i} (1 - S(t_i))^{R_i} \prod_{i=1}^{n} (S(t_i') - S(t_i)) \tag{1}
\]

where \( R \) is the set of right-censored observations, \( L \) the set of left-censored observations, and \( I \) is the set of interval-censored observations. Eq. 1 shows how each type of censoring contributes differently to the likelihood function.

In the presence of interval-censored data, the nonparametric estimation of the survival function is obtained maximizing the likelihood function (Eq. 1) using the Turnbull estimator (Turnbull 1976). The Turnbull estimator considers right- and left-censored data as particular cases of interval-censored data, being the upper limit of the observation interval equal to infinity and the lower limit equal to zero, respectively.

Parametric models and ML estimates. If we can assume an appropriate distribution for the data, the use of parametric models furnishes more precise estimates of the survival function and other quantities of interest than nonparametric estimators. Usually, survival times are not normally distributed; instead, their distribution is often right skewed. Often, a log-linear model is chosen:

\[
Y = \ln(T) = \mu + \sigma W
\]

where \( W \) is the error term distribution. That is, instead of the survival time \( T \), its logarithmic transformation is modeled. In Klein and Moeschberger (1997) or Lindsay (1998), different possible distributions for \( T \) are presented, for example the log-normal or the Weibull distribution. In case of the former, \( W \) is the standard normal distribution; in case of the Weibull distribution, \( W \) is the smallest extreme value distribution.

If the log-normal distribution is chosen for \( T \), the survival function is given by:

\[
S(t) = 1 - \Phi \left( \frac{\ln(t) - \mu}{\sigma} \right)
\]

where \( \Phi(\cdot) \) is the standard normal cumulative distribution function, and \( \mu \) and \( \sigma \) are the model’s parameters.

If the Weibull distribution is chosen, the survival function is given by:

\[
S(t) = S_{\text{weib}} \left( \frac{\ln(t) - \mu}{\sigma} \right)
\]

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

The parameters of the log-linear model are obtained by maximizing the likelihood function (Eq. 1). The likelihood function is a mathematical expression that describes the joint probability of obtaining the data actually observed on the subjects in the study as a function of the unknown parameters of the model being considered. To estimate \( \mu \) and \( \sigma \) for the log-normal or the Weibull distribution, we maximize the likelihood function by substituting \( S(t) \) in Eq. 1 by the expressions given in Eq. 2 or 3, respectively.

Once the likelihood function is formed for a given model, specialized software can be used to estimate the parameters (\( \mu \) and \( \sigma \)) that maximize the likelihood function for the given experimental data. The maximization is obtained by numerically solving the following system of equations using methods like the Newton-Raphson method (Gómez 2002):

\[
\frac{\partial \ln L(\mu, \sigma)}{\partial \mu} = 0
\]

\[
\frac{\partial \ln L(\mu, \sigma)}{\partial \sigma} = 0
\]

For more details on likelihood functions see Klein and Moeschberger (1997), Lindsay (1998), or Gómez and others (2001).

In survival analysis, the mean time to failure (\( E(t) \)) is defined as:

\[
E(t) = \int_0^\infty S(\tau) d\tau
\]

In sensory shelf-life studies, \( E(t) \) would represent the mean storage time in which consumers would reject the product. For the log-normal and Weibull distributions, \( E(t) \) is given by:

\[
E(t) = \exp(\mu + \frac{1}{2} \sigma^2)
\]

\[
E(t) = \exp(\mu) \Gamma(1 + \sigma)
\]

respectively, where \( \Gamma(\cdot) \) denotes the Gamma function.

Materials and Methods

Samples

A commercial whole fat, stirred, strawberry-flavored yogurt with strawberry pulp was used. Pots (150 g) were bought from a local distributor, all from the same batch. The pots were kept at 4 °C, and some of them were periodically placed in a 42 °C oven, so as to cover the following accelerated storage times: 0, 4, 8, 12, 24, 36, and 48 h. These times were chosen because a preliminary experiment showed that the flavor deteriorated quickly up to approximately 12 h and then slowed down. Once samples had reached the storage time at 42 °C, they were refrigerated at 4 °C, until they were tasted; this refrigerated storage lasted between 1 and 3 d. Previous microbiological analysis (aerobic mesophiles, coliforms, yeasts, and molds) showed that the samples were fit for consumption. The Ethical Committee of our institute decided that all samples were adequate for tests on humans in the quantities to be served.

Consumer study

Fifty subjects who consumed stirred yogurt at least once a week were recruited from the town of Nueva de Julio (Buenos Aires, Argentina). They were presented with the 7 yogurt samples (0, 4, 8, 12, 24, 36, and 48 h storage time at 42 °C) monadically in random order. Fifty g of each sample was presented in a 70-ml plastic cup. Time between each sample was approximately 1 min. Water was available for rinsing. For each sample, subjects tasted the sample and answered the question: “Would you normally consume this product? Yes or No?” It was explained that this meant: If they bought the product to eat it, or it was served to them at their homes, would they consume it or not? The tests were conducted in a sensory laboratory with individual booths with artificial daylight type illumination,
temperature control (between 22 and 24 °C), and air circulation.

Calculations
The KaplanMeier and CensorReg procedures from S-PLUS (Insightful Corp., Seattle, Wash., U.S.A.) were used to estimate the survival function nonparametrically and the parameters $\mu$ and $\sigma$, respectively.

Results and Discussion

Raw data and censoring considerations
Table 1 presents the data for 5 of the 50 subjects to illustrate the interpretation given to each subject’s data.
Subject 1 was as expected in a shelf-life study; that is, he/she accepted the samples up to a certain storage time and then consistently rejected them. The data are interval censored because we do not know at exactly what storage time between 12 and 24 h the consumer would start rejecting the product. Twenty-two subjects presented this type of data.

Subject 2 accepted all samples. Supposedly, at a sufficiently long storage time ($T > 48$ h), the sample would be rejected; thus, the data are right censored. Eight subjects presented this type of data.

Subject 3 was rather inconsistent, rejecting the sample with 8 h storage, accepting at 12 h, and rejecting from 24 h onwards. Censoring could be interpreted in different ways. One possibility, shown in Table 1, is interval censoring between 4 and 24 h. Another possibility would be to consider the data as interval censored between 4 and 8 h (ignoring the subject’s answers after the 1st time he/she rejected the yogurt). In the present study we considered the data to be interval censored between 4 and 24 h. Eleven subjects presented this type of data.

Subject 4 was also rather inconsistent with alternating no’s and yes’s. His/her data were considered as left censored. Left censoring can be considered as a special case of interval censoring, with the lower bound equal to time = 0 (Meeker and Escobar 1998). But as the literature and statistical software distinguish it, we have also done so. The left censoring could be considered as $T \leq 4$ h or $T \leq 24$ h; the latter was taken in the present study. Five subjects were left censored.

Subject 5 rejected the fresh sample; either: (a) he/she was recruited by mistake, that is they did not like yogurt; (b) they preferred the stored product to the fresh product; or (c) they did not understand the task. It would not be reasonable to consider the results of these subjects in establishing the shelf life of a product. For example, a company would have to produce a yogurt with a different flavor profile for consumers who preferred the stored to the fresh product, and not encourage these consumers to consume an aged product. Four subjects presented this behavior of rejecting the fresh sample, and their results were not considered. In a study on sensory shelf life of sunflower oil, Ramírez and others (2001) found that 9 out of 60 consumers preferred the stored to the fresh product, and their results were not included in the sensory failure calculations.

Failure probability calculations
To date, there are no statistical tests to compare the goodness-of-fit of different parametric models used for interval-censored data. Therefore, visual assessment of how parametric models adjust to the nonparametric estimation is the common practice in choosing the most adequate model. Figure 1 shows how 6 standard distributions were fitted to the yogurt data. Details about each one of these distributions can be found in the literature (Meeker and Escobar 1998; Klein and Moeschberger 1997). Two of these distributions were finally chosen: the lognormal because of the good fit to the data, and the Weibull because of its previous use in food shelf-life modeling. The maximum likelihood estimates of the parameters of these models (with their standard errors in parentheses) were as follows (See Eq. 2 and 3):

- Lognormal: $\mu = 2.99 (0.15), \sigma = 0.93 (0.13)$
- Weibull: $\mu = 3.39 (0.13), \sigma = 0.79 (0.12)$

Figure 2 shows the estimation of the survival function using the lognormal model and the nonparametric estimation. For ex-

---

Table 1—Acceptance/rejection data for 5 subjects who tasted yogurt samples with different storage times at 42 °C

<table>
<thead>
<tr>
<th>Subject</th>
<th>Storage time (h)</th>
<th>Censoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 4 8 12 24 36 48</td>
<td>Interval: 12-24</td>
</tr>
<tr>
<td>2</td>
<td>yes yes yes yes yes yes yes</td>
<td>Right: &gt;48</td>
</tr>
<tr>
<td>3</td>
<td>yes yes no yes no no no</td>
<td>Interval: 4-24</td>
</tr>
<tr>
<td>4</td>
<td>yes no yes yes no no no</td>
<td>Left: ≤ 24</td>
</tr>
<tr>
<td>5</td>
<td>no no yes yes yes yes no</td>
<td>Not consider</td>
</tr>
</tbody>
</table>

---

**Figure 1**—Probability of consumer rejecting the yogurt compared with storage time for 6 distribution models
ample, for a 50% probability of consumer acceptance adopted by some authors (Cardelli and Labuza 2001), the shelf life would be:

**Log-normal model:** 19.9 h, with 95% confidence bands of 14.9 and 26.6 h. These confidence bands are relatively wide, reflecting the uncertainty inherent to the censored data.

**Nonparametric estimation:** It lies between 12 and 24 h, since at 12 h of storage time, 66.7% of the consumers still accept the yogurt, whereas at 24 h 42% do.

Another value that could be of interest in defining shelf life is the mean time to failure, which can be obtained from the estimated models parameters as defined in Eq. 4 and 5. For the present data, these values (with 95% confidence bands in brackets) were:

**Lognormal distribution:**
\[ E(t) = 30.5 \text{ h} \ (21.1, 44.0) \]

**Weibull distribution:**
\[ E(t) = 27.6 \text{ h} \ (21.3, 35.7) \]

As the distributions are right skewed, these mean values were significantly larger than the median values: 19.9 h and 22.3 h for the lognormal and Weibull distributions, respectively.

**Conclusions**

To determine the sensory shelf life of food products, the focus has been set on the probability of a consumer accepting a product after a certain storage time. The survival analysis concept of censoring has been introduced, and it has been shown that different types of censoring have to be considered in a shelf-life study. Specialized software was used to fit experimental data to nonparametric and parametric models. An important aspect of this methodology is that experimental sensory work is relatively simple. In this work, 50 consumers each tasted 7 yogurt samples with different storage ages, answering “yes” or “no” to whether they would consume the samples. This information was sufficient to model the probability of consumers accepting the products with different storage times; from the models, shelf-life estimations were made. Future research in the application of survival analysis statistics to shelf-life studies should cover: recommendations of number of consumers and storage times necessary for desired statistical significance and power; accelerated storage studies where variables, such as storage temperature or humidity, would be considered as covariates; and how covariates related to the consumers, such as age or gender, influence shelf-life estimations.

**References**


Financial assistance from the Agencia Nacional de Promoción Científica (PICT-99) and the Ciencia y Tecnología para el Desarrollo Program (Project XI.16).

Authors Hough and Curia are with the Comisión de Investigaciones Científicas de PCIA, Buenos Aires, Argentina. Authors Langohr and Gómez are with the Dept. d’Estadística i Investigació Operativa, Unitat Politécnica de Catalunya, Pau Gargallo 5, 08028-Barcelona, Spain. Direct inquiries to Guillermo Hough (E-mail: guille@hough.cyt.edu.ar).