SORT 37 (2) July-December 2013, 175-188

Quantile estimation of the rejection distribution of food products integrating assessor values and interval-censored consumer data

Klaus Langohr[1], Guadalupe Gómez[1] and Guillermo Hough[2]

Abstract

Fitting parametric survival models with interval-censored data is a common task in survival analysis and implemented in many statistical software packages. Here, we present a novel approach to fit such models if the values on the scale of interest are measured with error. Random effects ANOVA models are used to account for the measurement errors and the likelihood function of the parametric survival model is maximized with numerical methods. An illustration is provided with a real data set on the rejection of yogurt as a function of its acid taste.

MSC: 62N99, 62F99.

Keywords: Interval-censored data; maximization of the likelihood function; parametric survival model; sensory shelf-life data.

1. Introduction

Since the publication of the work of Hough et al. (2003), survival data methods have become a common tool for the analysis of sensory shelf-life data of foods; see applications, among others, in Curia et al. (2005), Araneda et al. (2008) and Østli et al. (2013). The methodology has also been applied to determine consumer acceptance limits of sensory defects (Hough et al., 2004), and to optimize the concentration of food ingredients (Garitta et al., 2006).

[1] Department of Statistics and Operations Research, Universitat Politècnica de Catalunya/BARCELONATECH, Spain. E-mail: klaus.langohr@upc.edu

[2] Departamento de Evaluación Sensorial de Alimentos, Instituto Superior Experimental de Tecnología Alimentaria, Nueve de Julio, Buenos Aires, Argentina.

Received: December 2012
Accepted: July 2013

A typical shelf-life study consists of storing food samples for different lengths of time. For each time, consumers evaluate the product and report their acceptance or rejection. For example, for a yogurt study (Curia et al., 2005), samples were stored for 0, 14, 28, 42, 56, 70, and 84 days. A typical consumer's response would be: accept, accept, accept, reject, reject, reject, and reject, for each one of the respective times. This consumer's data is interval-censored between 28 and 42 days. Another consumer might accept all samples, and in this case the data would be right-censored at 84 days. A left-censored consumer would be one who rejects the sample which has only been stored for 14 days. Thus, data on the acceptance or rejection of a food product are generally interval-censored – including both left and right censoring as particular cases – where the intervals contain the real unknown values of rejection on the scale of interest; for the yogurt example the scale of interest was storage time.

The methodology proposed by Hough et al. (2003) furnishes the estimation of the rejection quantiles of interest for a given parametric model such as the Weibull, loglogistic or lognormal distribution. It assumes that the endpoints of the observed censoring intervals are all measured exactly without any error. Another instance of the application of this methodology is found in Sosa et al. (2008) who estimated the optimum concentration of salt in French-type bread from a consumer's perspective. They prepared samples of bread with 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, and 4.2 g sodium chloride per 100 g of flour. Since the weighing error of these salt quantities could be considered negligible, the values could be taken as exact.

However, the values of the independent variable may not always be free of error. Consider the case of a yogurt manufacturer who has applied survival analysis methodology to establish sensory shelf life of his product as described by Hough et al. (2003). If this manufacturer, in the future, wants to test a formulation change and make sure the sensory shelf life is still valid, he/she would have to assemble approximately 100 consumers (Hough et al., 2007). This is a costly and time-consuming experiment. If the critical descriptor (Hough, 2010) of yogurt from a consumer's perspective is acid taste, it would be of interest to the manufacturer to know how much the acid taste can increase before reaching 50% consumer rejection. If this acid taste cut-off value is known, then for future shelf-life determinations of the yogurt, the manufacturer can assemble a trained panel to measure acid taste instead of assembling the costly consumer panel. In this case the independent variable of the survival analysis experiment would become acid taste. These values are measured on a sensory scale by a sensory panel consisting of trained assessors. Presented with the same stimulus (a sample of yogurt) different assessors can produce different responses on the sensory scale; and the same assessor can produce different responses to sample replicates; thus the measurements are with error.

The objective of this work is to estimate the quantiles of the rejection distribution of a given food product integrating data from trained assessors and from consumers. Trained assessors provide the value of a certain characteristic of the product, such as the acid taste of yogurt. These values are random and subject to two sources of variability, one inherent to the assessor and the other corresponding to the specific acid taste of the product. Consumers will evaluate the same products as the trained assessors, stating their acceptance or rejection. Consumer data is interval-censored, where the endpoints of the intervals are random variables corresponding to the trained panel's measurements. Unlike other works, the law of the censoring endpoints is taken into account.

The content of the remainder of this work is the following. After describing the data of both trained assessors in Section 2 and consumers in Subsection 3.1, the likelihood function of the model and data under study is derived in Subsection 3.2. In Section 4, we give details on how to maximize this function in the framework of a parametric model and how to estimate the parameters and quantiles of interest. Section 5 presents the application of the estimation proposal to the motivating data set under study and in Section 6 the main results of this work are discussed.

2. Trained assessors: data, model, and analysis

For the sake of a better understanding, throughout the following sections, we use the data on the rejection of yogurt as a function of its acid taste. Yogurt samples were stored different times so that they would develop different levels of acid taste. These samples were given both to a panel of trained assessors and to consumers. Assessors received three replicates of each sample and measured their level of acid taste on a common scale from 0 to 100. Consumers received a single replicate of each sample and judged whether or not they would accept it.

2.1. Data and Model

A panel of J trained assessors are given K replicates of I different samples of yogurt which correspond to I different degrees of acid taste. Acid taste, denoted by X_{ijk}, was measured on a sensory scale from 0 (minimum acid taste) to 100 (maximum value), where k stands for replication ($k = 1,\dots,K$), j for assessor ($j = 1,\dots,J$), and i for sample ($i = 1,\dots,I$). In our motivating example, we have $K = 3$, $J = 13$, and $I = 6$. A graphical representation of all trained assessors' data is shown in Figure 1.

It is assumed that the data of a given sample i, $i = 1,\dots,I$, come from a one-way random effects ANOVA model:

$$X_{ijk} = \mu_i + \alpha_{ij} + \epsilon_{ijk}, \tag{1}$$

where $\alpha_{ij} \sim \mathcal{N}(0,\sigma^2_{b;i})$ and $\epsilon_{ijk} \sim \mathcal{N}(0,\sigma^2_{w;i})$. For sample i, the grand mean μ_i, representing the unknown acid taste of sample i, is the parameter of interest, α_{ij} is the random effect corresponding to assessor j, $j = 1,\dots,J$, and $\sigma^2_{b;i}$ and $\sigma^2_{w;i}$ denote, respectively, the between and within-assessors variances. Note that $\sigma^2_{b;i}$ is equivalent to the covariance between two observations on the same assessor (Vittinghoff et al., 2005, Chap. 8).

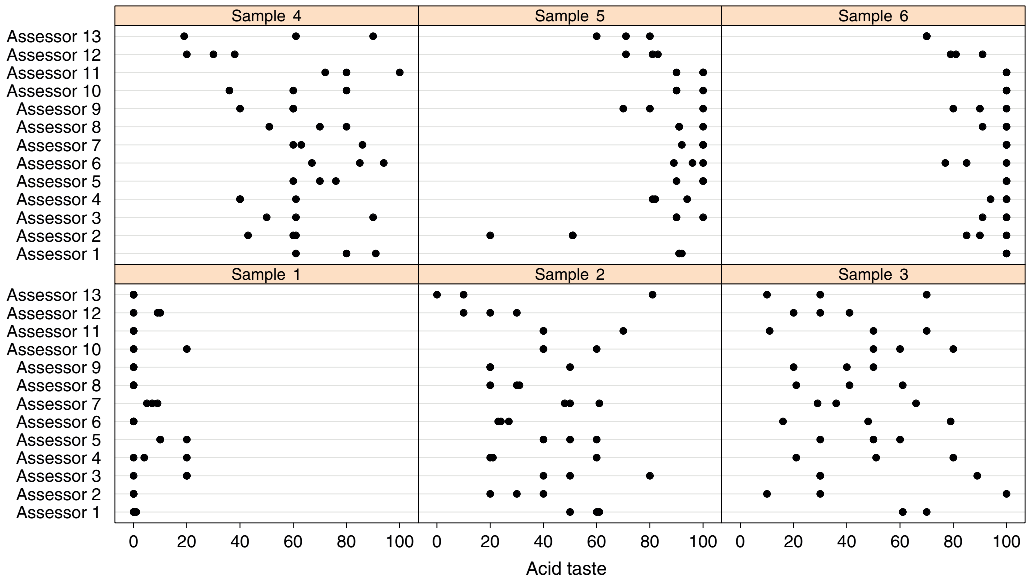

Figure 1: *Estimates of acid taste of yogurt given by 13 trained assessors on three replications of six different samples. Acid taste was measured on a 0 (minimum) to 100 (maximum) sensory scale.*

The model assumes independence among the assessors and between α_{ij} and ϵ_{ijk}. Hence, the overall variance of X_{ijk} is the sum of both variance components, that is, $\mathrm{Var}(X_{ijk}) = \sigma^2_{b;i} + \sigma^2_{w;i}$. In addition, and without loss of generality,

$$0 < \mu_1 < \cdots < \mu_I < 100, \tag{2}$$

where 100 may be substituted by any other value determined to be the maximum of the scale of interest.

2.2. Estimation

The estimator of the grand mean μ_i, $i = 1, \ldots, I$, is given by the overall mean of all $J \cdot K$ measurements given for each sample:

$$\hat{\mu}_i = \bar{X}_i = \frac{1}{J}\frac{1}{K}\sum_{j=1}^{J}\sum_{k=1}^{K} X_{ijk},$$

and its variance is equal to

$$\mathrm{Var}(\bar{X}_i) = \frac{1}{J \cdot K}(\sigma^2_{w;i} + K \cdot \sigma^2_{b;i}).$$

See Appendix A for further details.

Given the normal distribution assumption in model (1), the distribution of $\bar{X}_i$ is

$$\bar{X}_i \sim \mathcal{N}\big(\mu_i, \frac{1}{J \cdot K}(\sigma^2_{w;i} + K \cdot \sigma^2_{b;i})\big), \tag{3}$$

and for $J = 13$ and $K = 3$, the overall mean $\bar{X}_i$ follows a normal distribution with mean μ_i and variance $\frac{1}{39}(\sigma^2_{w;i} + 3 \cdot \sigma^2_{b;i})$.

Several estimators exist for both variance components including the restricted maximum likelihood estimators shown in (4). They are based on the between and within-assessors sum of squares ($\mathrm{SS}_{b;i}$ and $\mathrm{SS}_{w;i}$):

$$\begin{aligned}
\hat{\sigma}^2_{w;i} &= \min\Big(\frac{\mathrm{SS}_{w;i}}{J(K-1)}, \frac{\mathrm{SS}_{w;i} + \mathrm{SS}_{b;i}}{J \cdot K - 1}\Big), \\
\hat{\sigma}^2_{b;i} &= \max\left(0, \frac{1}{K}\left(\frac{\mathrm{SS}_{b;i}}{J-1} - \frac{\mathrm{SS}_{w;i}}{J(K-1)}\right)\right),
\end{aligned} \tag{4}$$

where $\mathrm{SS}_{w;i} = \sum_{j=1}^{J} \sum_{k=1}^{K} (X_{ijk} - \bar{X}_{ij})^2$ and $\mathrm{SS}_{b;i} = \sum_{j=1}^{J} K \cdot (\bar{X}_{ij} - \bar{X}_i)^2$. Herein, $\bar{X}_{ij}$ is the mean of assessor j's values for the ith sample. For a detailed discussion on these and other possible estimators, see Chapter 2 in Sahai and Ojeda (2004). In Appendix B, we give some details on computational aspects with R (The R Foundation for Statistical Computing).

Applying the previous formulas to our data set, we obtain sample mean estimates, the between and within-assessors standard deviations as well as the standard errors of $\bar{X}_i$ for all six samples which are shown in Table 1. We observe, for example, that the within-assessors standard deviations for samples 1 and 6 are much smaller then the rest; this is also reflected in Figure 1. When assessors measure samples with very low (sample 1) or very high (sample 6) acidities, they are all in agreement as to how to score these extreme samples. However, when intermediate acidities (samples 2 to 5) are presented, assessors can differ in their scores due to different perceptions and responses. This can be observed in the case of sample 3, where the estimated between-assessors variance is virtually 0, indicating that the variability observed in the estimation of the acid taste of this sample can be attributed entirely to the within-assessors variance.

Table 1: *Estimation results for model* (1) *for all six samples.*

	Sample					
	1	**2**	**3**	**4**	**5**	**6**
$\hat{\mu}_i$	4.2	39.2	46.2	62.7	85.8	93.4
$\hat{\sigma}_{w;i}$	5.9	17.7	23.4	17	8.6	5.6
$\hat{\sigma}_{b;i}$	3.9	8.5	0.0	11.4	15	8.3
$\hat{\sigma}_{\bar{X}_i}$	1.4	3.7	3.7	4.2	4.4	2.5

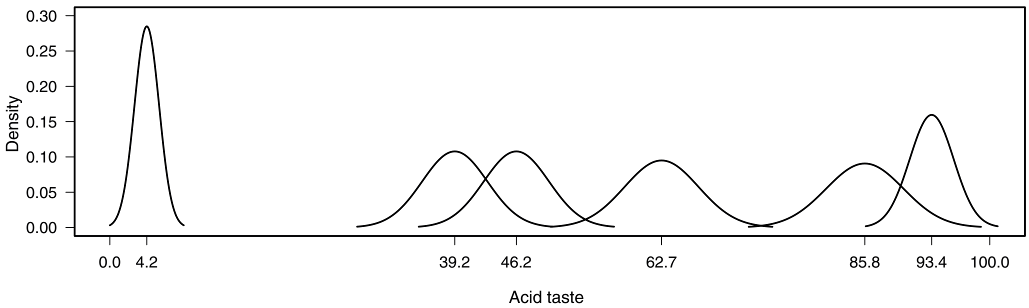

Figure 2: *Density functions of sample mean estimators* $\bar{X}_i$, $i = 1, \ldots, 6$.

In Figure 2, the density functions of all six mean estimators are represented assuming $\sigma^2_{b;i}$ and $\sigma^2_{w;i}$ are equal to the estimates shown in Table 1. In the following section, we will show how the uncertainty in the estimation of μ_i is taken into account in the analysis of the consumers' data.

3. Consumers: Data, rationale, and likelihood function

3.1. Data, rationale, and notation

In Section 1, the typical characteristics of a shelf-life study were presented. It was mentioned that the resulting data from the consumers, who are given the food product under study, are generally interval-censored containing the unknown value of rejection. Note that survival analysis methods can be applied to any positive random variable, for instance, yogurt's acid taste, as it is applied in the study that motivated the present work.

A total of $n = 74$ subjects are presented with $I = 6$ yogurt samples of different acid taste in a random order and have to answer the question whether they would normally consume such a yogurt or not. Based on their answers (acceptance/rejection), intervals of degrees of acid taste are determined that contain the acid taste from which a yogurt would be rejected. The interval for subject m, $m = 1, \ldots, n$, is of either of the two following types, where l_m and r_m indicate the sample number: $(l_m, r_m]$ or, in case of a right-censored observation, (l_m, ∞). In case of a left-censored observation, we define $l_m = 0$. Hence, $l_m \in \{0, \ldots, I\}, \forall m$, and $r_m \in \{1, \ldots, I, \infty\}, \forall m$. We denote the corresponding (unknown) acid tastes on the sensory scale from 0 to 100 by $(X_{l_m}, X_{r_m}]$, $m = 1, \ldots, n$.

In Table 2, the frequency distribution of the intervals obtained is shown. It can be seen that there are no left-censored and two right-censored data. That is, all subjects accepted sample 1, and two subjects did not reject any of the six samples. The fact that apart from the two right-censored observations not all of the remaining intervals are of type $(l_m, l_m + 1]$ is due to certain inconsistencies of the consumers' answers such as a sequence of "accept, reject, accept, reject". In that particular case, the interval obtained

Table 2: *Frequency distribution of intervals that contain rejection value.*

Interval	***n***	**%**
$(1,2]$	5	6.8
$(1,4]$	5	6.8
$(1,5]$	7	9.4
$(2,3]$	10	13.5
$(2,5]$	7	9.4
$(3,4]$	12	16.2
$(3,6]$	1	1.4
$(4,5]$	25	33.8
$(6,\infty)$	2	2.7
Total	74	100

is of type $(l_m, l_m+3]$, for example $(1,4]$; see Hough et al. (2003) for a more detailed discussion.

If the sample numbers were substituted by the corresponding estimated acid tastes shown in Table 1 without taking into account the uncertainty of the estimation, one could apply standard nonparametric methodology such as the Turnbull estimator (Turnbull, 1976) to estimate the quantiles of interest. The resulting graphical representation is shown in Figure 3 indicating, for example, that, according to this estimation, the median lies between 46.2 and 62.7.

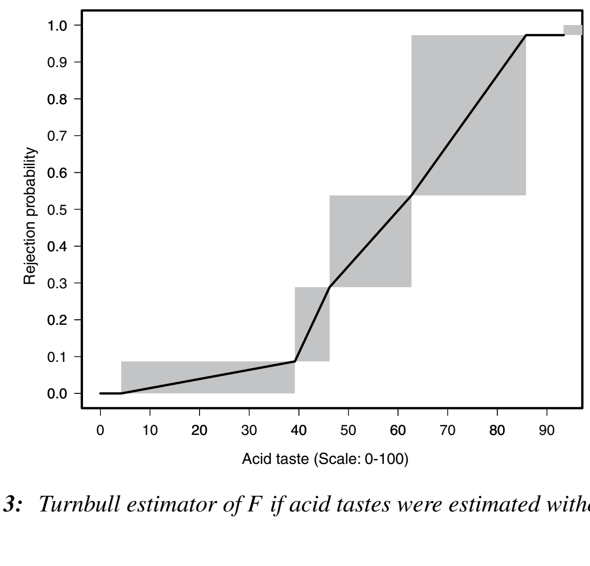

Figure 3: *Turnbull estimator of F if acid tastes were estimated without error.*

3.2. The likelihood function

In the following, we denote the distribution function of the random variable T, the acid taste from which yogurts are rejected, by F_T.

Assuming non-informative censoring (Oller et al., 2004) and if the acid tastes were measured without error, the contribution to the likelihood function of subject m, whose rejection value lies in interval $(x_{l_m}, x_{r_m}]$, would be (Gómez et al., 2009)

$$L_m = F_T(x_{r_m}) - F_T(x_{l_m}). \tag{5}$$

However, the exact acid tastes are unknown and estimates obtained from the panel of the 13 trained assessors are given instead. For this reason, we substitute the unknown acid tastes by these estimates and account for the corresponding uncertainty by integrating over the whole range of $\hat{\mu}_i$, $i = 1, \ldots, I$, which are all real-valued numbers in $[0, 100]$ restricted to $x_{l_m} < x_{r_m}$. Hence, the likelihood contribution in (5) converts into

$$L_m = \int_0^{100} \int_0^{r} \big(F_T(r) - F_T(l)\big)\, dF_{\bar{X}_{l_m}}(l)\, dF_{\bar{X}_{r_m}}(r). \tag{6}$$

Given a sample of size n, $(l_m, r_m]$, $m = 1, \ldots, n$, and assuming independence among the observations, the likelihood function is

$$L = \prod_{m=1}^{n} \int_0^{100} \int_0^{r} \big(F_T(r) - F_T(l)\big)\, dF_{\bar{X}_{l_m}}(l)\, dF_{\bar{X}_{r_m}}(r). \tag{7}$$

In case of left and right-censored observations, that is $l_m = 0$ and $r_m = \infty$, respectively, the likelihood contribution in (6) reduces to the following respective single integrals: $L_m = \int_0^{100} F_T(r)\, dF_{\bar{X}_{r_m}}(r)$ (left censoring) and $L_m = \int_0^{100} \big(1 - F_T(l)\big)\, dF_{\bar{X}_{l_m}}(l)$ (right censoring).

4. Maximization of the log-likelihood function

To maximize the logarithm of the likelihood function (7), following Wang (2010), discrete supports for $\bar{X}_i$, $i = 1, \ldots, I$, with corresponding probability masses have to be chosen. We denote these by

$$S_i = \{s_{i_1}, \ldots, s_{i_{p_i}}\} \quad \text{and} \quad \mathbf{\Pi}_i = \{\pi_{i_1}, \ldots, \pi_{i_{p_i}}\},\ i = 1, \ldots, I, \tag{8}$$

respectively. Different discrete supports of $\bar{X}_i$ can be thought of. For example, using the notation in (8), the first and last element of each support could be:

- $s_{i_1} = 0$ and $s_{i_{p_i}} = 100$,
- $s_{i_1} = \bar{x}_{i-1}$ and $s_{i_{p_i}} = \bar{x}_{i+1}$ with $\bar{x}_0 = 0$ and $\bar{x}_{I+1} = 100$,
- $s_{i_1} = \max(0, \bar{x}_i - p \cdot \hat{\sigma}_{\bar{x}_i})$ and $s_{i_{p_i}} = \min(100, \bar{x}_i + p \cdot \hat{\sigma}_{\bar{x}_i})$ for some $p \in \mathbb{N}$.

In either case, the mesh size h should be kept constant over the whole support, choosing, for example, $h = 0.1$ or $h = 0.5$.

The resulting expression of the log-likelihood function for the likelihood function given in (7) is as follows:

$$l = \sum_{m=1}^{n} \ln\Big(\sum_{v=1}^{p_{r_m}} \sum_{w=1}^{p_{l_m}} \big(F_T(s_{r_{mv}}) - F_T(s_{l_{mw}})\big)\pi_{l_{mw}}\pi_{r_{mv}}\mathbb{1}\{s_{l_{mw}} < s_{r_{mv}}\}\Big), \tag{9}$$

where both indices, v and w, cover the ranges of the corresponding supports but are restricted to $s_{l_{mw}} < s_{r_{mv}}, \forall v, w$, because of (2).

Given that $\bar{X}_i$ follows a normal distribution according to (3) and defining $\sum_{\Pi_i} = \sum_{l=1}^{p_i} f_{\bar{X}_i}(s_{i_l})$, we propose the following probability masses $\mathbf{\Pi}_i$, which are proportional to the density function of $\bar{X}_i$ evaluated in each point of the support S_i:

$$\pi_{i_v} = f_{\bar{X}_i}(s_{i_v}) \Big/ \sum\nolimits_{\Pi_i},\ v = 1, \dots, p_i,$$

where

$$f_{\bar{X}_i}(x) = \frac{1}{\sqrt{2\pi}\hat{\sigma}_{\bar{X}_i}} \exp\left(-\frac{1}{2}(\frac{x - \bar{x}_i}{\hat{\sigma}_{\bar{X}_i}})^2\right).$$

Hence, the expression of the log-likelihood function (9) becomes:

$$\begin{aligned} l = & \sum_{m=1}^{n} \ln\Bigg(\Big(\sum_{v=1}^{p_{r_m}} \sum_{w=1}^{p_{l_m}} \big(F_T(s_{r_{mv}}) - F_T(s_{l_{mw}})\big) \\ & \cdot \frac{1}{2\pi\hat{\sigma}_{\bar{X}_{l_m}}\hat{\sigma}_{\bar{X}_{r_m}}} \exp\Big(-\frac{1}{2}\big((\frac{s_{l_{mw}} - \bar{x}_{l_m}}{\hat{\sigma}_{\bar{X}_{l_m}}})^2 + (\frac{s_{r_{mv}} - \bar{x}_{r_m}}{\hat{\sigma}_{\bar{X}_{r_m}}})^2\big)\Big)\mathbb{1}\{s_{l_{mw}} < s_{r_{mv}}\}\Big) \Big/ \sum\nolimits_{\Pi_{l_m}} \cdot \sum\nolimits_{\Pi_{r_m}}\Bigg) \\ = & \sum_{m=1}^{n} \Bigg(\ln\Big(\sum_{v=1}^{p_{r_m}} \sum_{w=1}^{p_{l_m}} \big(F_T(s_{r_{mv}}) - F_T(s_{l_{mw}})\big) \\ & \cdot \frac{1}{2\pi\hat{\sigma}_{\bar{X}_{l_m}}\hat{\sigma}_{\bar{X}_{r_m}}} \exp\Big(-\frac{1}{2}\big((\frac{s_{l_{mw}} - \bar{x}_{l_m}}{\hat{\sigma}_{\bar{X}_{l_m}}})^2 + (\frac{s_{r_{mv}} - \bar{x}_{r_m}}{\hat{\sigma}_{\bar{X}_{r_m}}})^2\big)\Big)\mathbb{1}\{s_{l_{mw}} < s_{r_{mv}}\}\Big) - \ln\Big(\underbrace{\sum\nolimits_{\Pi_{l_m}} \cdot \sum\nolimits_{\Pi_{r_m}}}_{\blacksquare}\Big)\Bigg). \end{aligned}$$

and since $\blacksquare$ does not depend on F, the log-likelihood function to be maximized is

$$\begin{aligned} l = & \sum_{m=1}^{n} \ln\Big(\sum_{v=1}^{p_{r_m}} \sum_{w=1}^{p_{l_m}} \big(F_T(s_{r_{mv}}) - F_T(s_{l_{mw}})\big) \\ & \cdot \frac{1}{2\pi\hat{\sigma}_{\bar{X}_{l_m}}\hat{\sigma}_{\bar{X}_{r_m}}} \exp\Big(-\frac{1}{2}\big(\frac{(s_{l_{mw}} - \bar{x}_{l_m})^2}{\hat{\sigma}^2_{\bar{X}_{l_m}}} + \frac{(s_{r_{mv}} - \bar{x}_{r_m})^2}{\hat{\sigma}^2_{\bar{X}_{r_m}}}\big)\Big)\mathbb{1}\{s_{l_{mw}} < s_{r_{mv}}\}\Big). \end{aligned} \tag{10}$$

In case of left and right-censored data, the contributions to the log-likelihood function are, respectively:

$$l_m = \ln\Big(\sum_{v=1}^{p_1} F_T(s_{1v})\frac{1}{\sqrt{2\pi}\hat{\sigma}_{\bar{X}_1}}\exp\big(-\frac{1}{2}(\frac{s_{1v}-\bar{x}_1}{\hat{\sigma}_{\bar{X}_1}})^2\big)\Big),$$

$$l_m = \ln\Big(\sum_{w=1}^{p_I} \big(1-F_T(s_{Iw})\big)\frac{1}{\sqrt{2\pi}\hat{\sigma}_{\bar{X}_I}}\exp\big(-\frac{1}{2}(\frac{s_{Iw}-\bar{x}_I}{\hat{\sigma}_{\bar{X}_I}})^2\big)\Big).$$

As pointed out in the introduction, our objective consists of estimating the quantiles of the rejection distribution under different parametric models. That is, we will substitute F by different expressions according to the parametric choices for T as shown in the following section.

5. Quantile estimation for parametric models

Three parametric laws, which are commonly used for shelf-life studies of foods (Hough et al., 2003), are considered for the random variable of interest T:

- Weibull with shape parameter k, scale parameter λ, distribution function given by $F_T(t) = 1-\exp(-(t/\lambda)^k)$, and $t_\alpha = \lambda \cdot \ln(\frac{1}{1-\alpha})^{1/k}$ as the quantile α,
- loglogistic with shape parameter k, scale parameter λ, $F_T(t) = 1-\frac{1}{1+(t/\lambda)^k}$, and $t_\alpha = \lambda(\frac{\alpha}{1-\alpha})^{1/k}$,
- lognormal with parameters μ and σ, $F_T(t) = \Phi\big(\frac{\ln(t)-\mu}{\sigma}\big)$, and $t_\alpha = \exp\big(\mu+\sigma\cdot\Phi^{-1}(\alpha)\big)$.

For sample i, $i = 1,\dots,I$, we have chosen a discrete support with first element given by $s_{i_1} = \max(0, \bar{x}_i - 3\cdot\hat{\sigma}_{\bar{x}_i})$, last element given by $s_{i_{p_i}} = \min(100, \bar{x}_i + 3\cdot\hat{\sigma}_{\bar{x}_i})$, and with mesh size equal to 0.1. These supports cover intervals on the domain of $\bar{X}_i$ of probability masses larger than 0.99 for each sample. With these choices, the computation time for the maximization of the log-likelihood function takes about 25 seconds with the Intel i7 processor (1.73 GHz) under Windows 7. Technical details on the implementation in R are given in Appendix B.

The maximization of function (10) yields the parameter estimates and five quantiles as shown in Table 3. Whereas the standard errors are returned together with the parameters' estimates, the delta method is used in order to compute the standard errors of the log-transformed quantiles. 95% confidence intervals are computed for $\ln(t_\alpha)$ and the exponential transformation is applied to obtain the confidence intervals for t_α. They are, hence, not symmetric with respect to $\hat{t}_\alpha$.

Table 3: *Estimates obtained under different parametric models: parameter estimates are shown together with standard errors, quantile estimates together with 95% confidence intervals.*

	Weibull	**Loglogistic**	**Lognormal**
$\hat{\boldsymbol{\theta}}$ $(s.e.(\hat{\boldsymbol{\theta}}))$	$\hat{k} = 4.113$ (0.467)	$\hat{k} = 6.510$ (0.805)	$\hat{\mu} = 4.044$ (0.036)
	$\hat{\lambda} = 65.138$ (2.292)	$\hat{\lambda} = 57.426$ (2.153)	$\hat{\sigma} = 0.263$ (0.029)
Quantiles (95%-CI)			
0.1	37.7 ([32.2, 44.2])	41.0 ([36.5, 46.1])	40.7 ([36.5, 45.4])
0.25 (Q1)	48.1 ([43.0, 53.8])	48.5 ([44.4, 53.0])	47.8 ([43.9, 52.1])
0.5 (Median)	59.6 ([55.1, 64.5])	57.4 ([53.3, 61.9])	57.0 ([53.0, 61.3])
0.75 (Q3)	70.5 ([65.9, 75.4])	68.0 ([62.6, 73.8])	68.1 ([63.0, 73.6])
0.9	79.8 ([74.3, 85.7])	80.5 ([72.4, 89.5])	79.9 ([72.5, 88.1])

We can see, for example, that the estimated median under the Weibull model is 59.6 and that the corresponding 95% confidence interval ranges from 55.1 to 64.5. That is, under the Weibull model, 50% of all consumers are expected to reject yogurt with an acid taste above 59.6 and this value would serve as the cut-off value for the yogurt manufacturer if the objective is to produce yogurt whose acid taste is rejected by at most 50% of all consumers. Note that the median estimates are somewhat lower in case of the two other parametric choices (57.4 and 57, respectively) and that all three estimates lie in the interval obtained by the nonparametric estimation shown in Figure 3.

6. Conclusions and discussion

In this work, we have presented an approach to fit parametric models to interval-censored data when the interval limits are not fixed values, but are rather measured with certain error. As stated in the introduction, survival analysis methodology has so far been used to estimate rejection probabilities in food products as a function of variables of interest such as storage time which were measured exactly. However, there are other situations in which the variable of interest is not error-free, such was the case of acid taste in yogurt presented as an example in this work. We have developed a model to take account of the variability in the measurement of the independent variable.

Since the maximization of the likelihood function with such data is not implemented in statistical software, we have accomplished the parameter estimation in R with different functions of contributed packages; see Appendix B. The R code used can be provided on request from the authors.

The results obtained permit us to draw conclusions about the rejection distribution of a given food product based on a scale whose values are estimated by a trained panel. However, from a statistical point of view, our primary interest is the comparison of the obtained results with the ones of the method that ignores the uncertainty of the sample mean estimation. It could be expected that our approach would yield larger standard errors and confidence intervals, nonetheless, the results (not shown here) are fairly similar. For example, the standard errors of both parameters of the Weibull distribution do only differ in the second decimal place among both methods, whereas they are even virtually the same considering the lognormal distribution. Therefore, the differences of the quantiles obtained with both methods as well as the widths of the corresponding confidence intervals are notably small. The same findings held when we used broader discrete supports for $\bar{X}_i$, $i = 1, \ldots, 6$.

Another approach to estimate the parameters is to use multiple imputations as described in Rubin (1987). For each of B runs, sample mean values would be generated from the normal distributions (3) and the parametric models would be fitted assuming these values were measured error-free. The parameters estimates are then obtained as means over the B estimates obtained. We did this for $B = 1000$ obtaining similar parameter estimates (results not shown) but with larger standard errors (between 18% and 44% larger) reflecting both sources of variances: between and within-imputation variances. We, therefore, do not recommend this approach.

Two aspects of interest, which were not addressed in this work, are the nonparametric estimation of F and methods to judge the goodness-of-fit of a given parametric choice. These are relevant topics for further research.

In summary, final results showed small differences in quantile estimations between our model and the ad hoc calculations that did not consider variability. Whether these small differences will hold for most practical applications is difficult to predict. Our recommendation is for researchers to apply the complete model presented in this work in order to be sure that their quantile estimations are correct.

Acknowledgements

The authors are grateful to the research group *Grup de Recerca en Anàlisi Estadística de la Supervivència* (GRASS) for the fruitful discussions. This research was partially supported by Grant MTM2012-38067-C02-01 of the Ministerio de Economía y Competitividad (Spain). We also thank the reviewers of this paper for their useful comments and suggestions resulting in an improvement of the manuscript.

A. Variance of the sample mean estimator $\bar{X}_i$

In the following, the variance of the sample mean estimators $\bar{X}_i$, $i = 1, \ldots, I$, is derived. Remember that independence is assumed among assessors and that the covariance between two observations on the same assessor is equal to the between-assessors variance:

$$\begin{aligned}
\mathrm{Var}(\bar{X}_i) &= (\frac{1}{J}\frac{1}{K})^2 \mathrm{Var}\big(\sum_{j=1}^{J}\sum_{k=1}^{K} X_{ijk}\big) = \frac{1}{(J\cdot K)^2}\sum_{j=1}^{J}\mathrm{Var}\big(\sum_{k=1}^{K} X_{ijk}\big) \\
&= \frac{1}{(J\cdot K)^2}\cdot J\cdot \mathrm{Var}\big(\sum_{k=1}^{K} X_{ijk}\big) \\
&= \frac{J}{(J\cdot K)^2}\Big(\sum_{k=1}^{K}\mathrm{Var}(X_{ijk}) + 2\cdot\sum_{k=2}^{K}\sum_{k^*=1}^{k-1}\mathrm{Cov}(X_{ijk}, X_{ijk^*})\Big) \\
&= \frac{J}{(J\cdot K)^2}\big(K\cdot(\sigma^2_{b;i}+\sigma^2_{w;i}) + 2\cdot\frac{1}{2}\cdot(K-1)\cdot K\cdot\sigma^2_{b;i}\big) \\
&= \frac{1}{J\cdot K}(\sigma^2_{w;i} + K\cdot\sigma^2_{b;i}).
\end{aligned}$$

B. Computational Issues

All computations of this work were carried out with R (The R Foundation for Statistical Computing), version 3.0.1. Following, we give some details on the functions used.

The estimates of the one-way random effects ANOVA model shown in Table 1 are obtained by fitting model (1) with function `lme` of package `nlme` (Pinheiro et al., 2013). This function, which uses the restricted maximum likelihood estimators in (4) for the variance components by default, could also handle unbalanced designs with different numbers of replicates among assessors.

The maximization of the log-likelihood function (10) under different parametric models was accomplished with function `mle2` of the contributed package `bbmle` (Bolker and R Development Core Team, 2012). This function returns both the maximum likelihood estimates and their standard errors. As initial values for the parameters to be estimated, which are required by the maximization algorithm, one can choose the parameter estimates that are obtained by fitting the corresponding parametric model under the assumption that sample means were measured error-free.

References

Araneda, M., Hough, G. and de Penna, E. (2008). Current-status survival analysis methodology applied to estimating sensory shelf life of ready-to-eat lettuce (*lactuta sativa*). *Journal of Sensory Studies*, 23(2), 162–170.

Bolker, B. and R Development Core Team. (2012). *bbmle: Tools for general maximum likelihood estimation*. R package version 1.0.5.2.

Curia, A., Aguerrido, M., Langohr, K. and Hough, G. (2005). Survival analysis applied to sensory shelf life of yogurts-I: Argentine formulations. *Journal of Food Science*, 70(7), S442–S445.

Garitta, L., Serrat, C., Hough, G. and Curia, A. (2006). Determination of optimum concentrations of a food ingredient using survival analysis statistics. *Journal of Food Science*, 71, S527–S531.

Gómez, G., Calle, M. L., Oller, R. and Langohr, K. (2009). Tutorial on methods for interval-censored data and their implementation in R. *Statistical Modelling*, 9(4), 259–297.

Hough, G. (2010). *Sensory Shelf Life Estimation of Food Products*. Boca Raton, Florida: CRC Press.

Hough, G., Calle, M. L., Serrat, C. and Curia, A. (2007). Number of consumers necessary for shelf-life estimations based on survival analysis statistics. *Food Quality and Preference*, 18, 771–775.

Hough, G., Garitta, L. and Sánchez, R. (2004). Determination of consumer acceptance limits to sensory defects using survival analysis. *Food Quality and Preference*, 15, 729–734.

Hough, G., Langohr, K., Gómez, G. and Curia, A. (2003). Survival analysis applied to sensory shelf life of foods. *Journal of Food Science*, 68(1), 359–363.

Oller, R., Gómez, G. and Calle, M. L. (2004). Interval censoring: model characterizations for the validity of the simplified likelihood. *The Canadian Journal of Statistics*, 32, 315–325.

Østli, J., Esaiassen, M., Garitta, L., Nøstvold, B. and Hough, G. (2013). How fresh is fresh? perceptions and experience when buying and consuming fresh cod fillets. *Food Quality and Preference*, 27(1), 26–34.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team. (2013). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1–109.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley and Sons.

Sahai, H. and Ojeda, M. (2004). *Analysis of Variance for Random Models 1: Balanced Data. Theory, Methods, Applications and Data Analysis*. Boston: Birkhäuser.

Sosa, M., Flores, A., Hough, G., Apro, N., Ferreyra, V. and Orbea, M. (2008). Optimum level of salt in french-type bread. Influence of income status, salt level in daily bread consumption and test location. *Journal of Food Science*, 73, S392–S397.

Turnbull, B. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B*, 38, 290–95.

Vittinghoff, E., Glidden, D., Shiboski, S. and McCulloch, C. (2005). *Regression Methods in Biostatistics*. Springer.

Wang, Y. (2010). Maximum likelihood computation for fitting semiparametric mixture models. *Statistics and Computing*, 20(1), 75–86.