On Estimating the True Score of a Krapfen

The DaisyBio Team[1] and Julian Owezarek[2,3]

[1] *Technische Universität München, School of Life Sciences*
[2] *Max-Planck-Institut für Plasmaphysik, Numerische Methoden der Plasmaphysik*
[3] *Technische Universität München, School of Computation, Information and Technology*

January 21, 2026

Abstract

During the Fasching season, Krapfen consumption increases dramatically, while reliable estimates of their taste quality remain scarce and largely anecdotal. Taste is a latent, subjective variable subject to considerable uncertainty, prior beliefs, and sampling noise. In this paper, we propose a Bayesian framework for statistically sound estimation of Krapfen taste.

Contents

1 Introduction **1**

2 Krapfen Score Estimation - no AI **2**
- 2.1 The ML estimate . . . 3
- 2.2 The a posteriori distribution . . . 3
- 2.3 Example . . . 4

1 Introduction

During the carnival season, known in many parts of the world—and especially in Austria and Bavaria—as Fasching, the consumption of Krapfen reaches levels that would alarm both cardiologists and classical statisticians alike. These jam-filled pastries appear ubiquitously in offices, bakeries, and social gatherings, often accompanied by strong opinions about their quality. Statements such as "This one is amazing" or "Last year's were better" are routinely made, typically without any formal uncertainty quantification. This lack of methodological rigor represents a significant gap in the current literature.

Estimating the true "taste" of a Krapfen is a nontrivial statistical problem. Taste is latent, subjective, noisy, and strongly influenced by prior beliefs (e.g., nostalgia, hunger level, or the baker's reputation). Furthermore, direct measurements are costly, as each observation requires the partial or complete destruction of the sample. In this paper, we argue that Bayesian inference provides a natural and deliciously appropriate framework for addressing these challenges. By explicitly modeling prior expectations, observational uncertainty, and posterior updates after each bite, we construct a statistically principled method for inferring the underlying taste quality of Krapfen during Fasching season.

2 Krapfen Score Estimation - no AI

Assumption 2.1. *We assume that it is possible to subjectively rate a Krapfen between* 0 *and* 10 *(integer values only) to ones own satisfaction, where a score (rating) of* 0 *means that the subjective taste of a Krapfen could not have been worse, and a score of* 10 *means that the subjective taste of a Krapfen could not have been better.*

Assumption 2.2. *We assume that the taste-score of every type of Krapfen (of every bakery) is subject to a Binomial distribution. That is to say that for every type of Krapfen $i \in$ {'Nuss-Nougat(Rischart)', 'Nuss-Nougat(Ziegler)', ..., 'Pistazie(Rischart)', ...} there exists a true value $p_i \in [0,1]$ such that the true taste-score s of a random sample of a Krapfen k in the set of all currently acquirable Krapfen of type i, write*

$$K_i = \{k \ : \ k \text{ is a Krapfen currently acquirable (in Munich) of type } i\},$$

is binomial distributed with parameters $p = p_i$ and $n = 10$, i.e.,

$$\begin{aligned} P(s) &= \\ &= P(\text{"the true taste-score of this } k \in K_i \text{ is } s \in 0, \dots, 10\text{"}) \\ &= \binom{10}{s} p_i^s (1-p_i)^{10-s}. \end{aligned}$$

Of course, p_i is a priori unknown.

{ass:3}

Assumption 2.3. *We assume that the rating that our experts give a Krapfen with true taste-score $s \in \{0, \dots, 10\}$ is again Binomial distributed, hence a truely good (or bad) Krapfen will more probably be given a good (or bad) taste-score by Judith and Alex.*

We propose a Bayesian ansatz for estimating the true taste-score of a Krapfen. That is, we estimate an a priori guess for p_i using data from a survey. Typically, it is difficult to conduct large surveys, at times one may only be able to ask two colleagues (Lisi, Quirin) for their opinion. There is nothing we can do. We will then compute the ML-estimate $\hat{p}_i$ of the value p_i using the survey data. Finally, we estimate the true value p_i by taking into account the taste-score of our Krapfen experts (Judith, Alex). Their subjective measured score will define an a posteri probability distribution, which as far as I know, should again be a Binomial distribution with parameter $p = \tilde{p}_i$. The parameter $\tilde{p}_i$, or equivalently the expected value of the a posteriori distribution $\mathbb{E} = 10\tilde{p}_i$ shall serve as the best possible estimate for the true quality of all Krapfen $k \in K_i$.

Even if the a posteriori distribution is not a Binomial distribution, its expected value shall serve as best possible estimate for the true quality of all Krapfen $k \in K_i$.

2.1 The ML estimate

Fortunately, the ML-estimate for a binomial distributed random variable is easy to calculate (assuming I can trust the weird YouTube video I watched): Given two survey score $s_1, s_2 \in \{0, \ldots, 10\}$ of Lisi and Quirin, the ML-estimate $\hat{p}_i$ is

$$\hat{p}_i = \frac{s_1 + s_2}{20}.$$

For example, let's assume Lisi and Quirin rate Krapfen of type $K_i =$ "Nuss-Nougat von Rischart" with $s_1 = 8$ and $s_2 = 6$ respectively. Then the a priori probability distribution is given by

$$P(s) = \binom{10}{s} 0.7^s 0.3^{10-s},$$

as $\hat{p}_i = (8+6)/20 = 0.7$. This approach can be generalized to larger surveys.

2.2 The a posteriori distribution

We wish to compute the a posteriori distribution $\tilde{P}$ defined by

$$\tilde{P}(s) = P(s|\text{Event})$$

where "Event" is the outcome of the i.i.d. random variable that is the score given by the experts. Applying Bayes formula yields

$$\begin{aligned} \tilde{P}(s) &= P(s \mid \text{Event}) \\ &= P(\text{Event} \mid s)\frac{P(s)}{P(\text{Event})}. \end{aligned}$$

Two issues arise when computing the r.h.s. of this equation. The first is posed by the term $P(\text{Event})$. We compute

$$\begin{aligned} P(\text{Event}) &= \sum_{i=0}^{10} P(\text{Event} \cap i) \\ &= \sum_{i=0}^{10} P(\text{Event} \mid i)P(i). \end{aligned}$$

The second is posed by the remaining conditional probabilities. The question is: Given a specific Krapfen with true taste-score s, what is the probability that Judith or Alex give this Krapfen a rating of $\tilde{s}$? According to Assumption 2.3, this conditional probability can be computed using a Binomial distribution with parameter $p = s/10$.

We are able to compute the a posteriori distribution. For simplicity, we summarize how to compute each relevant quantity. For that matter, we denote the scores of Judith and Alex with $\tilde{s}_1$ and $\tilde{s}_2$ respectively.

$$\begin{aligned} P(s) &= \binom{10}{s} \hat{p}_i^s (1-\hat{p}_i)^{10-s} \\ P(\text{Event} \mid s) &= P(\text{Judith rates } \tilde{s}_1 \mid s) P(\text{Alex rates } \tilde{s}_2 \mid s) \\ P(\text{Person rates } \tilde{s} \mid s) &= \binom{10}{\tilde{s}} (s/10)^{\tilde{s}} (1-(s/10))^{10-\tilde{s}} \\ P(\text{Event}) &= \sum_{i=0}^{10} P(\text{Event} \mid i) P(i) = \sum_{i=0}^{10} P(\text{Judith rates } \tilde{s}_1 \mid i) P(\text{Alex rates } \tilde{s}_2 \mid i) P(i) \end{aligned}$$

Finally we conclude

$$\begin{aligned} \tilde{P}(s) &= P(\text{Event} \mid s) \frac{P(s)}{P(\text{Event})} && (1) \\ &= \binom{10}{\tilde{s}_1} (s/10)^{\tilde{s}_1} (1-(s/10))^{10-\tilde{s}_1} \cdot \binom{10}{\tilde{s}_2} (s/10)^{\tilde{s}_2} (1-(s/10))^{10-\tilde{s}_2} && (2) \\ &\cdot \frac{\binom{10}{s} \hat{p}_i^s (1-\hat{p}_i)^{10-s}}{\sum_{i=0}^{10} \left[\binom{10}{\tilde{s}_1} (i/10)^{\tilde{s}_1} (1-(i/10))^{10-\tilde{s}_1} \binom{10}{\tilde{s}_2} (i/10)^{\tilde{s}_2} (1-(i/10))^{10-\tilde{s}_2} \binom{10}{i} \hat{p}_i^i (1-\hat{p}_i)^{10-i} \right]} && (3) \end{aligned}$$

We note that "Event" and the taste-score s of the Krapfen tasted are not independent, hence we can't write $P(\text{Event} \mid s) = P(\text{Event})P(s)$!

The expected value $\mathbb{E}$ of $\tilde{P}$, the best possible estimate for the true quality of all Krapfen $k \in K_i$, can then readily be computed as

$$\mathbb{E} = \sum_{s=0}^{10} s \tilde{P}(s).$$

2.3 Example

We illustrate our approach using an example. Let i be "Nuss-Nougat_Rischart". A sufficiently large group of 5 sane people would expect to rate a random Krapfen in K_i on

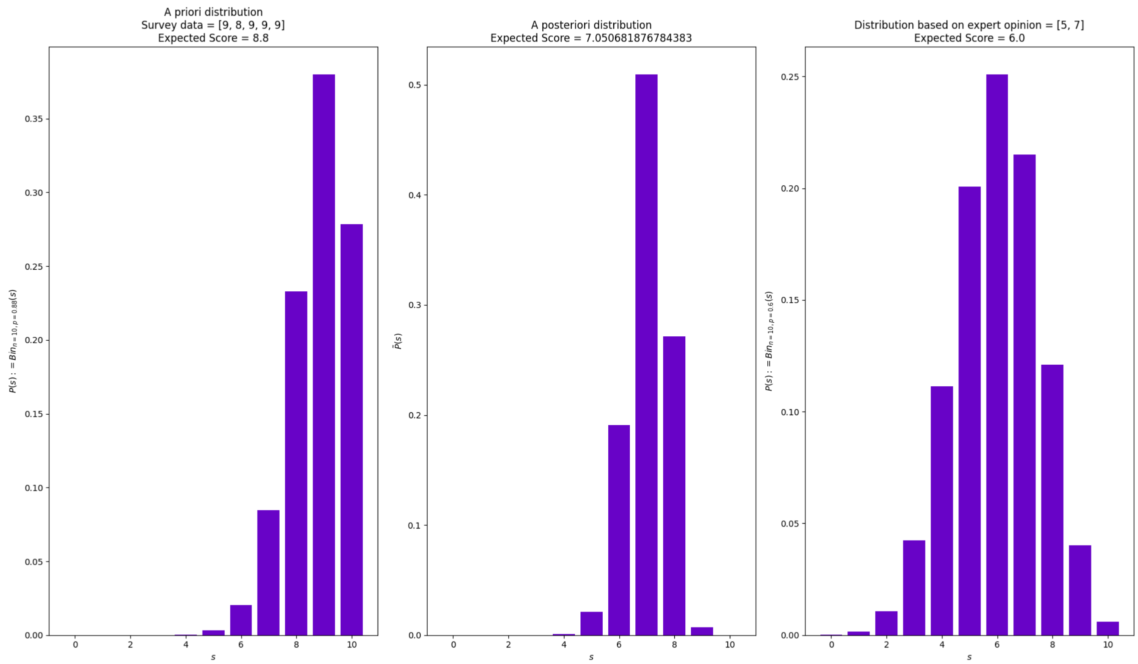

Figure 1: An illustration of the example.

avarage 8.8. This avarage value determines the ML-estimate $\hat{p}_i = 0.88$, which then defines the a priori Binomial distribution with parameters $n = 10, p = 0.88$.

Our experts rate a sample Krapfen $k \in K_i$ 5 and 7 (they are a haters). Not taking into account the prior, they would come up with a shared score of 6, the mean of a Binomial distribution with parameters $n = 10, p = (5 + 7)/20$.

Taking into account the prior, we obtain a better estimate for the true score, in this case ≈ 7.05.