

In statistical inference, the role of estimators is to recover unknown parameters from observed data, and statisticians leverage decision theory to evaluate these estimators by comparing their risk; however, like most problems in statistics, there is debate on whether a Bayesian or frequentist approach is superior. This essay advocates for Christian Robert's view, asserting that a well-chosen prior not only encapsulates Nature's inherent uncertainty and allows the Bayes risk to be, at worst, equal to and often lower than the minimax risk found in frequentist methods. To support Robert's claim, we begin by outlining the core principles of discrete statistical inference concerning decision theory, including estimators, parameters, expected loss, and minimax risk. We then introduce the Bayesian framework by detailing the role of priors and integrated risk, culminating in a formal proof that the Bayes risk does not exceed the minimax limit. Finally, we defend against Philip Stark's critique regarding the subjectivity of priors by demonstrating that some degree of subjectivity is unavoidable in decision theory.

We begin by defining some of the fundamental elements of statistical decision theory. To introduce the nomenclature, we let  $\mathcal{X}$  be the observation space, representing all possible data you might observe. Further, we let  $\Theta$  be the parameter space, capturing all conceivable "states of nature." Finally, we let  $\mathcal{A}$  be the action space, consisting of all actions or estimators you can choose; we allow all of these spaces to be finite and discrete for simplicity. For this paper, an action  $a \in \mathcal{A}$  generally implies estimating  $\theta \in \Theta$  with a certain formula  $\alpha(x)$  known as the *estimator*. The observations  $x \in \mathcal{X}$  are connected to the parameter  $\theta \in \Theta$  by the probability mass function  $p(x|\theta)$ , referred to as the data-generating process (DGP) [1]. In a discrete setting, the DGP describes the probability of an observation  $x \in \mathcal{X}$  under a given parameter  $\theta$ . The primary objective of statistical inference is to infer underlying properties of the DGP and  $\theta$  [2]. To illustrate the connection to decision theory, suppose you flip a coin  $n$  times and wish to estimate the parameter  $\theta$  corresponding to the true proportion of times your coin lands heads. In this case, the DGP  $p(x|\theta)$  defines a process where, for each coin flip, there is a probability  $\theta$  of the coin landing heads and a probability  $1 - \theta$  of landing tails.

One action  $a_1 \in \mathcal{A}$  is to propose the ‘sample mean’ estimator  $\alpha_1(x) = \frac{1}{n} \sum_{i=1}^n x_i$ , whereas another action  $a_2 \in \mathcal{A}$  is to naively propose  $\alpha_2(x) = 1$  (every flip is heads). It can be shown<sup>1</sup> that  $a_1$  proposes an estimator which maximizes the likelihood of the observed data under the DGP [3], whereas  $a_2$ ’s estimator is biased, thus trivially  $a_1 \succ a_2$ . Unless necessarily distinct, we henceforth use estimators  $\alpha$  and the actions  $a$  proposing them interchangeably.

To quantify the preference orderings beyond the simple heuristics mentioned in the coin-flipping case, statisticians leverage loss functions [4], which we denote  $\mathcal{L}(\theta, \alpha)$ . The loss function represents the error associated with proposing a “bad” estimation of the  $\theta$  of interest. Thus, the best evaluation of this function is a zero loss; therefore,  $\mathcal{L}(\theta, \alpha) \geq 0$  [5]. From a decision-theoretic perspective, the objective of the decision-maker is to propose an estimator  $\alpha$  which minimizes this loss. Since the actual value of parameter  $\theta$  is often unknown, statisticians base their ordering of estimators on the *expected* loss. However, precisely how we define *expected* relies upon whether one takes a frequentist or Bayesian approach.

Under the frequentist paradigm, the data  $x \in \mathcal{X}$  are considered random because they arise from repeated sampling via the DGP  $p(x|\theta)$ . Meanwhile,  $\theta$  is treated as a fixed but unknown constant in the parameter space  $\Theta$ . In the coin-flipping example, a frequentist would assume that the coin has a fixed but unknown probability  $\theta$  of landing heads, and thus  $p(x|\theta)$  governs each flip outcome. Hence, to evaluate a proposed estimator  $\alpha$ , the frequentist approach focuses on expected loss, similar to how Peterson [6] considers the expected utility. Specifically, we define the expected loss (EL) as the product of the probability of observing  $x \in \mathcal{X}$  and the loss associated with estimating  $\theta$  with  $\alpha(x)$ ,

$$\text{EL}(\theta, \alpha) = \mathbb{E}_\theta[\mathcal{L}(\theta, \alpha)] = \sum_{x \in \mathcal{X}} \mathcal{L}(\theta, \alpha(x)) p(x|\theta) \quad (1)$$

We also refer to the above as a risk function [7]. From this definition of expected loss, we introduce the concept of “minimax” through a game-theoretic analogy of a game against

---

<sup>1</sup>Given  $p(x_{1:n}|\theta) = \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i}$ , the log-likelihood of the  $n$  observations is  $\ell(x, \theta) = \log(\theta) \sum_{i=1}^n x_i + \log(1-\theta) \sum_{i=1}^n (1-x_i)$ . Maximizing with respect to  $\theta$  yields  $\hat{\theta}_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^n x_i = \alpha_1(x)$ .

Nature. In this framework, we aim to select an estimator  $\alpha \in \mathcal{A}$  that *minimizes* our expected loss. Meanwhile, Nature acts as our adversary, selecting a parameter  $\theta \in \Theta$  (i.e., a “state of the world”) in an attempt to *maximize* our expected loss [8]. The risk in such a game is known as the “minimax risk”, which we define as

$$r^* = \min_{\alpha \in \mathcal{A}} \left\{ \max_{\theta \in \Theta} \left\{ \text{EL}(\theta, \alpha) \right\} \right\} \quad (2)$$

While the minimax risk  $r^*$  is occasionally criticized as being overly conservative [6], the ability of an estimator to be the best in the worst case scenario (which we refer to as the “minimax guarantee”) is desirable for many real-world applications including management of financial portfolios [9].

Having defined the minimax risk in Equation (2) and the corresponding ‘worst-case’ guarantee, we now turn to *The Bayesian Choice* [5], where Christian Robert shows that Bayesian decision theory yields a risk that is *no greater* than the minimax risk—and often strictly better. We first introduce the notion of a *prior* to explain the Bayesian paradigm. In a discrete parameter space, the prior is a function  $\pi$  which maps every  $\theta \in \Theta$  to a probability. By definition, the prior satisfies  $\sum_{\theta \in \Theta} \pi(\theta) = 1$  where  $\pi(\theta)$  is the probability that  $\theta$  is the “true” state of the world. In other words, a Bayesian would say that “ $\theta$  follows a  $\pi$ -distribution *a priori*.” In this framework, the data  $x \in \mathcal{X}$  still arise from the DGP  $p(x|\theta)$  (now known as the “likelihood”) but are treated as *fixed* once observed. Bayesian methods instead place uncertainty in  $\theta$ , initially via  $\pi(\theta)$  and later in  $\pi(\theta|x)$ , the *a posteriori* distribution of  $\theta$  after observing  $x$ . In contrast, frequentist methods conceptualize  $x \in \mathcal{X}$  as potentially variable under repeated sampling, while  $\theta$  is fixed but unknown.

To evaluate a proposed estimator  $\alpha \in \mathcal{A}$ , the Bayesian approach focuses on the posterior expected loss (PEL). The PEL averages the loss  $\mathcal{L}(\theta, \alpha(x))$  across all possible values of  $\theta \in \Theta$  and is weighted by the posterior probability of the parameter  $\pi(\theta|x)$  conditioned on the observed value  $x$ .

The PEL is given by:

$$\text{PEL}(\pi, \alpha|x) = \mathbb{E}_\pi[\mathcal{L}(\theta, \alpha)|x] = \sum_{\theta \in \Theta} \mathcal{L}(\theta, \alpha(x))\pi(\theta|x) \quad (3)$$

The equation above considers the weighted loss across all  $\theta \in \Theta$  for a singular  $x$ , whereas Equation 1 weighs across all  $x \in \mathcal{X}$  for a singular  $\theta$ ; thus, the two measures are not commensurable. To bridge the gap between the Bayesian and frequentist paradigms, Robert introduces the notion of Integrated Risk and Bayes Risk<sup>2</sup>. The Integrated Risk is the frequentist risk averaged over the values of  $\theta$  according to their prior distribution  $\pi(\theta)$ . Integrated risk is useful for comparison because it associates a real number with every estimator, while the posterior expected loss varies with each observation  $x$ . Formally, we write the integrated risk as:

$$r(\pi, \alpha) = \sum_{\theta \in \Theta} \text{EL}(\theta, \alpha)\pi(\theta) = \sum_{\theta \in \Theta} \sum_{x \in \mathcal{X}} \mathcal{L}(\theta, \alpha(x))p(x | \theta)\pi(\theta) \quad (4)$$

The above formulation allows for comparison with the frequentist paradigm. To the skeptical reader, it may not seem immediately obvious how weighing the frequentist loss according to the prior  $\pi(\theta)$  fully captures the Bayesian approach. However, Robert proves that selecting an estimator  $\alpha \in \mathcal{A}$  which minimizes the purely-Bayesian posterior expected loss (see Equation 3) *also* minimizes the integrated risk [5]. Thus, despite conditioning on observed data, he concludes that the Bayesian approach is still justified by frequentist standards since it yields estimators that minimize the overall integrated risk. To argue in favour of a Bayesian approach over a frequentist one, Robert analyzes the best estimator (and its risk) under Equation 4, showing it yields risk less than or equal to the minimax standard. This ‘best estimator’ is a *Bayes estimator*. For a given prior distribution  $\pi$  and a loss function  $\mathcal{L}$ , the Bayes estimator  $\alpha^\pi(x)$  minimizes the posterior expected loss for every observation  $x \in \mathcal{X}$ . Consequently, it minimizes the overall integrated risk. From a constructive standpoint, the Bayes estimator can be defined by minimizing Equation 3 for each observation, i.e.

---

<sup>2</sup>Note that formal definitions of Bayes Risk date as far back as the 1980s with works from James O. Berger[10]. Robert himself cites these works as part of his argument.

$$\forall x \in \mathcal{X}, \alpha^\pi(x) = \operatorname{argmin}_{\alpha \in \mathcal{A}} \left[ \sum_{\theta \in \Theta} \mathcal{L}(\theta, \alpha(x)) \pi(\theta|x) \right]$$

Because  $\alpha^\pi$  minimizes the point-wise PEL for each  $x$ , it follows that  $r(\pi, \alpha^\pi)$  is also the minimal integrated risk for all  $\alpha \in \mathcal{A}$ . This minimum value is known as the *Bayes risk*.

We now turn to the relationship between Bayes risk and the frequentist minimax risk, a central element of Robert's argument. Specifically, we will formalize the mathematical justification for preferring the Bayesian paradigm; namely, that for every prior  $\pi \in \Pi$ , where  $\Pi$  is the set of priors on  $\theta$ , the Bayes Risk is *at most* equal to the frequentist minimax risk. We prove this in the finite case.

**Box 1.** Fix  $\pi \in \Pi$  and  $\alpha \in \mathcal{A}$ . Let  $\mathcal{X}$ ,  $\Theta$ , and  $\text{EL}(\theta, \alpha)$  be as stated. Assume for contradiction that the Bayes risk is larger than minimax, i.e.

$$\min_{\alpha \in \mathcal{A}} \left\{ \sum_{\theta \in \Theta} \text{EL}(\theta, \alpha) \pi(\theta) \right\} > \min_{\alpha \in \mathcal{A}} \left\{ \max_{\theta \in \Theta} \left\{ \text{EL}(\theta, \alpha) \right\} \right\}$$

Noticing that  $\alpha \in \mathcal{A}$  is fixed and  $\sum_{\theta \in \Theta} \pi(\theta) = 1$  by definition, we write that

$$\sum_{\theta \in \Theta} \text{EL}(\theta, \alpha) \pi(\theta) > \max_{\theta \in \Theta} \left\{ \text{EL}(\theta, \alpha) \right\} \cdot \sum_{\theta \in \Theta} \pi(\theta)$$

Consider  $\Theta' = \{\theta \in \Theta \mid \pi(\theta) > 0\} \subseteq \Theta$ . As the maximum expected loss right-hand side is constant, by summation laws we rewrite the above with respect to  $\Theta'$  as

$$\sum_{\theta \in \Theta'} \pi(\theta) \left( \text{EL}(\theta, \alpha) - \max_{\theta \in \Theta} \left\{ \text{EL}(\theta, \alpha) \right\} \right) > 0$$

However, since  $\pi(\theta) > 0$  and  $\Theta' \subseteq \Theta$  this implies the existence of a  $\theta \in \Theta$  for which the expected loss is greater than the maximal expected loss, which is a contradiction. Because this argument holds for any fixed  $\alpha$ , taking the minimum over  $\mathcal{A}$  preserves the inequality. Thus, the Bayes Risk is at most the maximin risk for any fixed  $\alpha \in \mathcal{A}$  and  $\pi \in \Pi$  as required.  $\square$

The minimaxity of Bayes risk follows directly from the above. Since the result holds for any fixed  $\pi$ , it also holds for the *least favourable* prior that maximizes the Bayes risk. In other words, the ‘worst-case’ Bayes risk is no greater than the minimax risk. Thus, the Bayesian procedure attains the ‘minimax guarantee’ in the worst case, just as the frequentist approach does. We now illustrate this minimax guarantee with a return to the coin-flipping example. We introduce two candidate priors that achieve the same ‘worst-case’ Bayes risk, matching the frequentist minimax risk.

**Box 2.** In addition to the previous setup where  $p(x|\theta) = \theta^x(1-\theta)^{1-x}$  and  $\alpha_1 = \frac{1}{n} \sum x_i$ , we have  $\alpha_2 = 0.25 + 0.5x$  as a candidate in  $\mathcal{A}$ . We further consider observations  $\mathcal{X} = \{1, 0\}$ , parameters  $\Theta = \{\frac{1}{2}, 0\}$  and loss function  $\mathcal{L}(\theta, \alpha) = |\theta - \alpha(x)|$ . We employ priors  $\pi_1 \sim \text{Discrete}\{(\theta_1, \theta_2), (\frac{1}{5}, \frac{4}{5})\}$  and  $\pi_2 \sim \text{Uniform}\{\theta_1, \theta_2\}$ . Using Equations 1-4, one can compute the following loss tables under frequentist expected loss and Bayes risk:

	$\alpha_1$	$\alpha_2$	$\max_{\Theta}$		$\pi_1$	$\pi_2$	$\min_{\mathcal{A}}$
$\theta_1$	0	0.25	0.25	$\alpha_1$	0.40	0.25	0.25
$\theta_2$	0.25	0.5	0.5	$\alpha_2$	0.25	0.25	0.25

From the above,  $r^* = \min_{\mathcal{A}}\{0.25, 0.50\} = 0.25$  and  $r(\alpha^\pi) = \max_{\Pi}\{0.25, 0.25\} = 0.25$ .

In the above example, both priors  $\pi_1$  and  $\pi_2$  are least favourable, as the Bayesian maximin risk is equivalent to the minimax risk. However, even small adjustments of the variables in Box 2 (i.e., slightly modifying  $\alpha_2$ ) restore strict preference of  $r(\alpha^\pi)$  over  $r^*$ .

We have thus explained and mathematically justified the central pieces of Christian Robert’s argument. By incorporating prior information via the integrated risk, one can often demonstrate that the Bayes risk is no greater than the frequentist  $r^*$  and thus shares the same ‘minimax guarantee.’ In other words, by leveraging prior information, the Bayesian approach performs at least as well (and often better than) the standard minimax procedure. However, this raises a concern about the subjectivity of  $\pi(\theta)$ , which we will explore in Philip

Stark's counterargument.

While Stark does not dispute the proof-theoretic result that Bayes risk cannot exceed minimax risk, he argues in *Constraints versus Priors* that introducing a prior distribution  $\pi$  adds extra information to the decision problem, which *artificially* lowers the Bayes risk relative to the minimax risk. Moreover, he criticizes the subjective nature of priors, arguing that the Bayesian approach places biased assumptions on the state of Nature [11].

To illustrate Stark's critique of incorporating extraneous information, we extend the coin-flipping example from Box 2 by introducing an additional observation space  $\mathcal{Y}$ . Specifically, we define  $y \in \mathcal{Y}$  as the outcome of a die rolled in a separate room. This event is unrelated to the coin flip, the parameter  $\theta$ , or the loss incurred by choosing estimator  $\alpha$ . Nevertheless, the uniform distribution  $p(y) = 1/6$  exists over  $\mathcal{Y}$ , satisfying  $\sum_{y \in \mathcal{Y}} p(y) = 1$  and  $p(y) \geq 0$  for  $y \in \{1, 2, \dots, 6\}$ . Consequently, we can adjust Equation 1 by marginalizing over this irrelevant observation space. The resulting *modified* expected loss (MEL) is defined as:

$$\text{MEL}(\theta, \alpha) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \mathcal{L}(\theta, \alpha(x)) p(x | \theta) p(y)$$

Notice that this construction mimics the definition of integrated risk in Equation 4, except rather than adding prior information  $\pi(\theta)$  we incorporate irrelevant  $p(y)$ . Indeed, one can show directly that  $\min_{\alpha \in \mathcal{A}} \text{MEL}(\theta, \alpha)$  is at most the minimax risk  $r^*$ , mirroring the relationship of Bayes risk to  $r^*$ . So the question arises: what makes the set of priors  $\Pi$  any different from purely extraneous information  $\mathcal{Y}$ ?

Stark argues that there is hardly any difference at all. He claims that by selecting a prior  $\pi \in \Pi$ , the statistician is assuming the state of Nature, which opens their methodology to subjectivity and human error [11]. Specifically, he argues that the Bayesian analyst “claim[s] to know more about how Nature generates the data” than the frequentist [11, p.8]. Recall our earlier game-theoretic description of the estimation loss as one in which our opponent, Nature, selects a parameter  $\theta \in \Theta$  in an attempt to maximize the loss  $\mathcal{L}$  of choosing estimator  $\alpha \in \mathcal{A}$  [8]. In the frequentist paradigm, the statistician does not know the rule by which Nature

selects  $\theta$  to generate the data  $\mathcal{X}$ . To Stark, the Bayesian framework is far more relaxed; a Bayesian assumes Nature has a distribution  $\pi$  it selects  $\theta$  from, which is shared *a priori* to both players. Hence, he insists that the Bayes risk approach, contrary to Robert’s claim of superiority, is, in fact, a strategy for an entirely different game and *cannot* be directly compared to the frequentist procedure. As he puts it, the reduced Bayes risk relative to minimax “is because the prior added information not present” in the frequentist constraints [11, p.10]. In other words, any sizable gap between Bayes and minimax risk should prompt us to question whether we genuinely share the analyst’s beliefs about how Nature selects  $\theta$  via  $\pi(\theta)$  or if it is potentially meaningless ‘extra’ information —akin to  $p(y)$  in MEL—that yields the difference. To summarize, Stark maintains that the inherent subjectivity in selecting prior alongside the risk of introducing irrelevant information undermines the Bayesian framework, rendering it no better than the frequentist approach.

To combat Stark’s objections, Robert would argue that when done prudently, the choice of a prior is not an arbitrary injection of extra information but a necessary aspect of statistical modeling. He would remind Stark that even in cases where the analyst lacks knowledge of the problem domain, they can employ an uninformative or “diffuse” prior distribution to objectively capture features of the parameter  $\theta$  without imposing their subjective beliefs [12]. By construction, the influence of a diffuse prior on the posterior is minimal [13]. Hence, the likelihood  $p(x|\theta)$  dominates the posterior, and the resulting Bayes estimator (and thus Bayes risk) depends on the observations  $x \in \mathcal{X}$  rather than the analyst’s subjective inputs. Importantly, such a diffuse prior is, at worst, least favourable; thus, the inequality proved in Box 1 holds without the injection of personal bias. Consequently, the choice of prior exerts a far more negligible impact on the integrated risk than Stark’s critique would imply.

Moreover, even if an analyst chooses a prior based on personal beliefs, subjectivity is an inherent aspect of decision-theoretic analysis. For example, as illustrated in Box 2, the loss function is arbitrary; one might just as easily opt for Huber loss [14], squared error loss [15], or many other alternatives instead of the absolute error. In choosing  $\mathcal{L}(\theta, \alpha) = |\theta - \alpha(x)|$



to evaluate the coin-flipping estimators, the analyst made an inherently subjective decision. Thus, to assess estimators, frequentist and Bayesian analysts make subjective decisions. Therefore, just as Stark challenges us to question *why* an analyst believes  $\pi$  accurately describes how Nature selects  $\theta$ , one might equally question whether  $\mathcal{L}$  truly captures the loss relationship between  $\theta$  and  $\alpha$  or even reflects a sensible model of Nature. In this sense, Robert would argue that Stark's argument is flawed, as he denounces the metaphysical assumptions of the Bayesian paradigm, which are equally present in the frequentist framework. Ultimately, he would argue, the Bayesian approach embraces the unavoidable subjectivity inherent in decision theory by selecting a prior that quantifies uncertainty about the natural world and, in doing so, often reduces the integrated risk.

In conclusion, we stand by the mathematically justified perspective of Christian Robert and other Bayesians, advocating for the Bayesian framework which not only meets the min-max guarantee but often exceeds it. Criticisms regarding the subjectivity of priors fall short when faced with the reality that all statistical methods rely on underlying assumptions that may introduce bias. Rather than undermining the Bayesian approach, these critiques reaffirm the importance of incorporating uncertainty in the state of Nature through prior distributions, reinforcing our thesis in favor of Bayesian methodology.

## References

- [1] Jun Tu and Guofu Zhou. Data-generating process uncertainty: What difference does it make in portfolio decisions? *Journal of Financial Economics*, 72(2):385–421, 2004.
- [2] Graham Upton and Ian Cook. *Oxford Dictionary of Statistics*. Oxford University Press, 2008.
- [3] Richard J. Rossi. *Mathematical Statistics: An Introduction to Likelihood Based Inference*. John Wiley & Sons, New York, 2018. p. 227.

- 
- [4] Abraham Wald. *Statistical Decision Functions*. Wiley, Oxford, England, 1950. Includes 76-item bibliography. PsycINFO Database Record (c) 2016.
- [5] Christian P. Robert. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer Texts in Statistics. Springer-Verlag New York, 2 edition, 2007.
- [6] Martin Peterson. *An Introduction to Decision Theory*. Cambridge University Press, New York, 2nd edition, 2017.
- [7] M. S. Nikulin. Risk of a statistical procedure. In *Encyclopedia of Mathematics*. EMS Press, 2001. Originally published in 1994.
- [8] V. Ulansky and A. Raza. Generalization of minimax and maximin criteria in a game against nature for the case of a partial a priori uncertainty. *Heliyon*, 7(7):e07498, Jul 2021.
- [9] Xiao-Tie Deng, Zhong-Fei Li, and Shou-Yang Wang. A minimax portfolio selection strategy with equilibrium. *European Journal of Operational Research*, 166(1):278–292, 2005. Metaheuristics and Worst-Case Guarantee Algorithms: Relations, Provable Properties and Applications.
- [10] James O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, 2nd edition, 1985.
- [11] Philip B. Stark. Constraints versus priors. Technical report, 2020. Department of Statistics, University of California, Berkeley.
- [12] Arnold Zellner. *Prior Distributions to Represent 'Knowing Little'*. John Wiley & Sons, New York, 1971.
- [13] Sarah Lee. A comprehensive guide to noninformative priors in bayesian analysis. *Number Analytics Blog*, March 2025.

- [14] Peter J. Huber. Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 53(1):73–101, 1964.
- [15] Trevor Hastie, Robert Tibshirani, and Jerome H. Friedman. *The Elements of Statistical Learning*. Springer, 2001.

## Appendix

Below is the code used to compute the loss tables in Box 2 and the comments following it. The interested reader may edit lines 3, 6, 10, 19, 26, 27, or 32 to see that the Bayes risk is always lower than the minimax risk except for the example given in Box 2.

```
1 # priors
2 pi_1 <- function(theta){
3     if(theta == 0.5){
4         return(0.2)
5     } else {
6         return(0.8)
7     }
8 }
9 pi_2 <- function(theta){
10     return(1/2)
11 }
12 Pi <- list(pi_1, pi_2)
13 # alphas
14 a_1 <- function(x){
15     return(mean(X))
16 }
17 a_2 <- function(x){
18     # a(0)=0.25, a(1)=0.75.
19     epsilon <- 0.00 # small change in epsilon -> bayes preferred
20     # 0.25 -> 0.25+(5*epsilon / 10) for freq
21     # 0.25 -> 0.25+(epsilon/10) for bayes
22     return(0.25 + (0.50 + epsilon)*x) # key piece for equality
23 }
24 A <- list(a_1, a_2)
```

```

25 # x, theta, PDF and loss
26 X <- c(0,1)
27 Theta <- c(1/2, 0)
28 pX <- function(x, theta){
29     return(dbinom(x, 1, theta))
30 }
31 L <- function(theta, a){
32     return(abs(theta - a))
33 }
34 # ----- FREQUENTIST RISK TABLE -----
35 frequentist_risk <- matrix(NA, nrow=length(Theta), ncol=length(A))
36 rownames(frequentist_risk) <- paste("theta_", Theta)
37 colnames(frequentist_risk) <- c("a1", "a2")
38 for (i in seq_along(Theta)) {
39     for (j in seq_along(A)) {
40         theta <- Theta[i]
41         risk <- 0
42         for (x in X) {
43             # expected loss (eq 1)
44             risk <- risk + L(theta, A[[j]](x)) * pX(x,
45                 theta)
46         }
47         frequentist_risk[i, j] <- risk
48     }
49 }
50 print("Frequentist_Risk_Table:")
51 print(frequentist_risk)
52 # ----- BAYES RISK TABLE -----

```

```

52 bayes_risk <- matrix(NA, nrow=length(Pi), ncol=length(A))
53 rownames(bayes_risk) <- c("pi_1", "pi_2", "uniform")
54 colnames(bayes_risk) <- c("a1", "a2")
55 for (k in seq_along(Pi)) {
56     prior <- Pi[[k]]
57     for (j in seq_along(A)) {
58         risk_bayes <- 0
59         for (i in seq_along(Theta)) {
60             # integrated risk, eq 4
61             risk_bayes <- risk_bayes + frequentist_risk[
62                 i, j] * prior(Theta[i])
63         }
64         bayes_risk[k, j] <- risk_bayes
65     }
66 }
67 print("Bayes_Risk_Table:")
68 print(bayes_risk)
69 # ----- COMPUTE -----
70 # minimax, eq 2
71 frequentist_minimax <- min(apply(frequentist_risk, 2, max))
72 print(paste("Frequentist_Minimax_Risk=", frequentist_minimax))
73 bayes_maximin <- max(apply(bayes_risk, 1, min))
74 print(paste("Bayesian_Maximin_(Bayes_Risk_under_least_favourable_
prior)=", bayes_maximin))

```