Assignment 4 = Exam 1: Measures and Probability
Due October 17, 2023

John McCuan

Problem 1 Let π be a probability measure on a measure space S.

(a) Given a subset $A \subset S$ with $\pi(A) > 0$, define the restriction probability measure ρ_A determined on the measure space S by the set A. Hint: Be sure to give a formula for $\rho_A(T)$ for a measurable set $T \subset S$.

(b) Use your definition from part **(a)** above to prove the following[1] **law of partition**:

If $A_1, A_2, \ldots, A_k$ are (pairwise) disjoint measurable sets in S with

$$\bigcup_{j=1}^{k} A_j = S \qquad \text{and} \qquad \pi(A_j) > 0 \qquad \text{for } j = 1, 2, \ldots, k,$$

and A is any measurable subset of S, then

$$\pi(A) = \sum_{j=1}^{k} \rho_{A_j}(A)\ \pi(A_j).$$

[1]This is a mathematical/measure theory version of what Orloff and Booth call the **law of total probability** in their Class 3 notes.

Problem 2 (independence and dependence) Let $\pi : \wp(S) \to [0,1]$ be a probability measure on a set S with finitely many elements. Given two sets A and B with $A, B \subset S$ we say A and B are **independent** if

$$\rho_B(A) = \pi(A) \tag{1}$$

where ρ_B is the probability restriction measure on S determined by B.

(a) Taking the heuristic idea that the probability of an outcome is the value of a probability measure on some set modeling the outcome, explain in words the condition (1) defining what it means for two sets to be independent in terms of probabilities involving outcomes modeled by the two sets:

> A and B are **independent** if the probability of the outcome modeled by the set A is the same as. . .

(b) Write down formulas for the values of the probability restriction measures

$$\rho_A : \wp(S) \to [0,1] \qquad \text{and} \qquad \rho_B : \wp(S) \to [0,1].$$

(c) Notice that the defining relation (1) for independence of sets is not symmetric with respect to the sets A and B.

(i) Prove that if (1) holds, then

$$\rho_A(B) = \pi(B) \tag{2}$$

where ρ_A is the restriction measure on S determined by A.

(ii) As in part **(a)** above, express in words and in terms of "probabilities" the meaning or interpretation of the condition (2).

(iii) Show (2) implies (1), so the two conditions are equivalent.

(d) Show the conditions (1) and (2) are equivalent to the symmetric condition

$$\pi(A \cap B) = \pi(A)\ \pi(B).$$

Problem 3 (Binomial measure) Let n be a natural number and let $S = \{0,1\}^n$. Recall that given p with $0 < p < 1$ fixed, the binomial measure $\beta : \wp(S) \to [0,1]$ is determined by

$$\beta(\{(\omega_1, \omega_2, \dots, \omega_n)\}) = p^{\#\{j\,:\,\omega_j=1\}}(1-p)^{\#\{j\,:\,\omega_j=0\}}.$$

(a) Express the binomial measure as a product measure in terms of the probability measure $\pi : \wp(\{0,1\}) \to [0,1]$ with $\pi(\{1\}) = p$.

(b) Taking $n = 3$ consider the sets

$$x^{-1}(\{j\}) \qquad \text{for } j = 0, 1, 2, 3,$$

where $x : S \to \mathbb{R}$ by

$$x(\omega_1, \omega_2, \omega_3) = \omega_1 + \omega_2 + \omega_3.$$

(i) Find $x^{-1}(\{j\})$ for $j = 0, 1, 2, 3$.

(ii) Taking $p = 1/2$, find $M(j) = \alpha(\{j\}) = \beta(x^{-1}(\{j\}))$ for $j = 0, 1, 2, 3$.

(iii) Taking $p = 3/4$, find $M(j) = \alpha(\{j\}) = \beta(x^{-1}(\{j\}))$ for $j = 0, 1, 2, 3$.

(c) Generalize/repeat part **(b)** for $n = 4, 5, 6$

(d) Compute the integral of x with respect to the binomial measure β (for general n and p).

Problem 4 (simple p-values) Consider the event

"flipping a coin n-times and recording the number of heads."

Let α denote the binomial induced measure with

$$\alpha(\{j\}) = \begin{pmatrix} n \\ j \end{pmatrix} p^j(1-p)^{n-j} \qquad \text{for } j = 0, 1, 2, \ldots, n. \tag{3}$$

Note: The value p appearing in (3) with $0 < p < 1$ is **not** the probability p featured in the name "p-values," but you should understand that probability by the end of this problem.

We take as a "null hypothesis" the statement

The coin used in the event above is a fair coin.

Complete parts **(a)-(d)** under the assumption that the null hypothesis holds, i.e., $p = 1/2$.

(a) Calculate the expectation x_* of $x : \{0,1\}^n \to \mathbb{R}$ by

$$x(\omega_1, \omega_2, \ldots, \omega_n) = \sum_{j=1}^{n} \omega_j$$

with respect to the Binomial measure $\beta : \wp(\{0,1\}^n) \to [0,1]$, i.e., calculate the integral

$$x_* = \int_S x$$

with respect to β where $S = \{0,1\}^n$.

(b) Calculate the expectation of the identity function on $\mathbb{R}$ with respect to the binomial induced measure $\alpha : \wp(\mathbb{R}) \to [0,1]$.

(c) Taking $n = 3$, consider the set

$$A = \{(\omega_1, \omega_2, \omega_3) \in S : x(\omega_1, \omega_2, \omega_3) \geq 2\}. \tag{4}$$

(i) What compound outcome does the set A model and what is the probabilistic interpretation of $\beta(A)$ in terms of the event "flipping a coin 3 times and recording the number of heads?"

(ii) Rewrite the set A in the form

$$A = \{(\omega_1, \omega_2, \omega_3) \in S : x(\omega_1, \omega_2, \omega_3) - x_* \geq \delta\}$$

for some $\delta > 0$.

(iii) What compound outcome does the set

$$B = \{(\omega_1, \omega_2, \omega_3) \in S : |x(\omega_1, \omega_2, \omega_3) - x_*| \geq \delta\}$$

model?

(iv) Find $\beta(B)$.

(d) Generalize/repeat part **(c)** for $n = 4, 5, 6$ replacing the relation

$$x(\omega_1, \omega_2, \omega_3) \geq 2$$

in (4) with

$$x(\omega_1, \omega_2, \ldots, \omega_n) \geq n - 1.$$

(e) The answers you got for $\beta(B)$ are not technically p-values. Technically, a p-value is both the value of a probability measure, i.e., a "probability," **and a statistic**. This means that technically you need a data set to get a p-value. The idea is that the existence of a certain data set may cast doubt on (or justify the rejection of) the null hypothesis.

(i) Say you actually flip a coin three times with the result "heads," "tails," "heads" corresponding to the model outcome $(1, 0, 1)$. Then the value you computed in part **(c)(iv)** is the p**-value** associated with the data from your event (or experiment). The fact that the value $\beta(B)$ depends on the data makes it a statistic.

How does $\beta(B)$ depend on the data? Why is it a statistic?

(ii) If the value of $\beta(B)$ is "high," then the idea is that the data gives you no reason to reject the null hypothesis: As far as this data goes, it may very well be the case that the coin is a fair coin. But if the value is "low," then perhaps the null hypothesis should be rejected.[2]

Should the p-value in this case be considered "high" or "low?"

(iii) Repeat part **(e)(ii)** for $n = 4, 5, 6$. For example, say you flip a coin four times and obtain an outcome involving $4 - 1 = 3$ heads. Find the p-value. Do you think it is "high," "low," or somewhere in between?

(iv) Given data $(\omega_1, \omega_2, \ldots, \omega_n) \in S$ corresponding to actual coin flipping for some general n, formulate the associated p-value determined by the data and the null hypothesis *the coin is fair*. Hint: If $x(\omega_1, \omega_2, \ldots, \omega_n) = k$, then the p-value is the probability, assuming the null hypothesis, that any data $a \in S$ collected in a similar manner, i.e., by flipping the coin n times, has $x(a)$ at least as far from the expected value x_* as the actual data $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$.

[2] If one rejects applied probability out of hand and in principle, then all null hypotheses are automatically rejected as well—without any p-test.

(f) Here is perhaps the most interesting part of this problem: The description of the calculation of a p-value defined/described in part **(e)** above, e.g.,

> *. . . the probability,* **assuming the null hypothesis***, . . .*

strongly suggests the calculation of the value of some restriction probability measure.

(i) What is the **domain** of the measure in question which is being restricted?

(ii) What collection of abstract outcomes does that domain model?

Trivia: The actual threshold according to which a p-value is considered "small" is fundamentally arbitrary (as far as I know). The value 0.05 however is the most commonly used threshold, and this apparently[3] goes back to the mathematician Pierre-Simone Laplace who first argued in 1837 that the calculation of a specific p-value somewhat greater than 0.05, namely 0.0897, could reasonably be taken to indicate his null hypothesis (in this case something to do with "random" selection of jury members for trials) was possibly correct while he interpreted another calculation of a p-value of 0.00468 to be reasonable evidence that the associated null hypothesis should be rejected.

It seems that Ronald Fisher picked up on Laplace's threshold and in his 1925 book *Statistical Methods for Research Workers* associated the specific threshold value 0.05 with the term "statistical significance." And this is the (arbitrary) value that has been more or less officially adopted ever since. It is amusing that this adoption has led to the assertion by famous statistician and executive director of the American Statistical Association (ASA) Ron Wasserstein that "statistical significance" has today become meaningless. Wasserstein's statement appeared in connection with the articles in a special issue of *The American Statistician* entitled "Statistical inference in the 21st century: a world beyond $p < 0.05$." Included was an official ASA statement warning[4] against the use of statistical significance and p-values including the following:

- Do not base your conclusions solely on whether an association or effect was found to be *statistically significant* (i.e., the p value passed some arbitrary threshold such as $p < 0.05$);
- Do not believe that an association or effect exists just because it was statistically significant;
- Do not believe that an association or effect is absent just because it was not statistically significant;
- Do not believe that your p value gives the probability that chance alone produced the observed association or effect or the probability that your test hypothesis is true;
- Do not conclude anything about scientific or practical importance based on statistical significance (or lack thereof).

[3]according to Justin Zeltzer `https://www.youtube.com/watch?v=4XfTpkGe1Kc`

[4]quoted from an article in favor of the use of p-values and the 0.05 threshold, "Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a concervative approach" in the journal *European Radiology Experimental.*