Assignment 5: Uncountable measure spaces
Due October 24, 2023

John McCuan

Problem 1 (Assignment 2 Problem 6 (review) expected value) Consider the six sided die with five sides marked “3” and one side marked “6” with outcomes modeled by the set

$$S = \{\text{one}, \text{two}, \text{three}, \text{four}, \text{five}, \text{six}\}$$

and the function $r : S \to \{3, 6\}$ by $r(\omega) = 3$, $\omega = \text{one}, \ldots, \text{five}$ and $r(\text{six}) = 6$.

(a) Find

$$\int_S r$$

with respect to the uniform probability measure on S.

(b) Find

$$\int_{\mathbb{R}} \text{id}_{\mathbb{R}} = \int_{\xi \in \mathbb{R}} \xi$$

with respect to the measure on $\mathbb{R}$ induced by the function r.

(c) Orloff and Booth in Example 1 of their Class 4 notes[1] on “Expected Value” call the number calculated above the **expected value** and ask “What would you **expect** the average of 6000 rolls to be?”

(i) Does the value they expect you to calculate depend on the number of rolls?

(ii) What is that value?

(iii) Explain why is it called the **expected value** (if you can).

[1]Note: These Class 4 notes come in two installments, and the one I'm referencing here is the second installment.

(d) Look carefully at the definition of **expected value** given at the bottom of page 1 of the notes of Orloff and Booth, and especially the expression for $E(X)$ at the top of the next page.

Orloff and Booth say they need a "discrete random variable" and the "probabilities" of its "values" for their definition. In terms of integration on a measure space with finitely many elements, identify the following:

(i) What corresponds to a "value" x_j?

(ii) What corresponds to the "probability" $p(x_j)$?

(iii) What is $E(X)$?

Problem 2 (geometric distribution, expected value, Orloff and Booth Class 4 notes, Example 9) Consider the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ as an adolescent measure space with measure $\gamma : \wp(\mathbb{N}) \to [0, 1]$ by

$$\gamma(\{j\}) = (1-p)^{j-1}p$$

where as usual p is a fixed number with $0 < p < 1$.

Event: (trials) Some auxiliary event has outcomes success and failure. We refer to the occurrence of this event as a "trial," and consider executing the event over and over again, i.e., executing trials, until a success is observed.

Outcome: The number of trails executed in order for a success to be observed.

Modeling/model sets: $S = \{0, 1\}$ models the outcomes, success and failure of the auxilliary event. The number of trials executed in order for a success to be observed is modeled by a natural number in $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Measure(s): Bernoulli measure $\beta : \wp(S) \to [0, 1]$ with $\beta(\{1\}) = p$ on a measure space $S = \{0, 1\}$ and the geometric probability measure $\gamma : \wp(\mathbb{N}) \to [0, 1]$.

(a) Calculate

$$\int_{\mathbb{N}} 1.$$

More properly, we should write something like $\int_{\mathbb{N}} \mathbb{1}$ where $\mathbb{1} : \mathbb{N} \to \mathbb{R}$ is the constant function with $\mathbb{1}(j) \equiv 1$, but I simply wrote the constant 1 as the integrand. (This is what is called a small "abuse of notation.")

(b) Complete the following steps to calculate

$$\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} = \int_{\xi \in \mathbb{R}} \xi$$

with respect to the measure on $\mathbb{R}$ induced by the function $x : \mathbb{N} \to \mathbb{R}$ by $x(j) = j$ for $j \in \mathbb{N}$.

(i) Show the value of the integral is given by the (Riemann) sum

$$\sum_{j=1}^{\infty} j(1-p)^{j-1} p. \tag{1}$$

(ii) Write the function value(s) appearing in the Riemann sum as

$$\mathrm{id}_{\mathbb{R}}(j) = j = \sum_{\ell=1}^{j} 1,$$

and draw an illustration of the ordered pairs (j, ℓ) involved in the resulting double sum.

(iii) Using your illustration as a guide, change the order of summation to write

$$\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} = p \sum_{\ell} \sum_{j} (1-p)^j. \tag{2}$$

(The important part here is to get the limits of summation correct.)

(iv) Use the formula

$$\sum_{k=0}^{\infty} \rho^k = \frac{1}{1-\rho} \tag{3}$$

for the sum of a geometric series[2] with ratio ρ satisfying $0 < \rho < 1$ to find the sum in (2).

(c) Express the integral in part **(b)** as an integral over $\mathbb{N}$ with respect to the geometric probability measure γ. (The main point of this part is the identification of the real valued function which is being integrated.)

---

[2] You'll need to use this formula basically twice, but actually infintely many times.

(d) Orloff and Booth in Example 9 of their Class 4 notes[3] on "Expected Value" calculate the average value of the identity function on a measure space $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$, i.e., the integral of the identity function, with respect to a (slightly different but equivalent) probability measure $\gamma_0 : \wp(\mathbb{N}_0) \to [0, 1]$. They call this the **mean of the geometric distribution**.

(i) Using the format above (Event, Outcomes, Modeling sets, Measures) describe the context of "trials" in which the measure γ_0 might be used/referenced.

(ii) Use the method of Orloff and Booth to calculate the integral in part **(b)** above in a different way:

- Consider the value of the integral you calculated in part **(a)** above as a function of p:

$$f(p) = \sum_{j=1}^{\infty} (1-p)^{j-1} p = 1.$$

  Assuming the sum can be differentiated termwise, take the derivative with respect to p. (You will need to use the product rule and the chain rule from calculus for this. Maybe you will need to look them up.)

- You should now be able to find a furmula for

$$\sum_{j=1}^{\infty} j(1-p)^{j-2} p.$$

  Multiply this formula by $1-p$ to get an expression for the integral/sum in (1) in terms of geometric series.

- Evaluate the geometric series using (3) to get the same answer you got in part **(b)** above.

[3] Note: These Class 4 notes come in two installments, and the one I'm referencing here is the second installment.

Problem 3 (geometric series) Follow these steps to prove (3).

(a) Take a **partial sum**

$$s_\ell = \sum_{k=0}^{\ell} \rho^k$$

and multiply by $(1-\rho)$. Cancel the terms that cancel.

(b) Use your formula $(1-\rho)s_\ell = \phi(\ell)$ from part **(a)** to solve for s_ℓ.

(c) Take the limit

$$\sum_{k=0}^{\infty} \rho^k = \lim_{\ell \nearrow \infty} s_\ell$$

to finish the proof.

Problem 4 (integral probability measure, Orloff and Booth Example 3, Class 5 notes (second installment)) Consider the PDF $\delta : [0,1] \to \mathbb{R}$ by

$$\delta(\omega) = c\omega^2$$

where c is some constant. Given that δ is the PDF of a probability measure $\pi : \mathcal{M} \to [0,1]$, determine the following:

(a) The value of the constant c.

(b) $\pi(\{\omega : \omega \le 12\})$.

(c) The CDF of π.

(d) The mean

$$\int_{\mathbb{R}} \omega$$

of π.

Problem 5 (uniform probability measures; variance) Consider the one parameter family of uniform probability measures $\upsilon = \upsilon_r : \mathcal{M} \to [0,1]$ with statistical range $[-r, r]$.

(a) Find the MDF of υ.

(b) Calculate $\upsilon([a,b])$ for every $a, b \in \mathbb{R}$ with $a < b$.

(c) Calculate the mean or expected value of υ.

(d) Calculate the variance σ^2 of υ and express the value σ^2 as a monotone function of the spread $2r$.

Problem 6 (integration for babies) Let $\alpha : \wp(\mathbb{R}) \to [0,1]$ be any generalized baby measure or any generalized adolescent measure on the real line. Observe that we can associate a PMF $M : \mathbb{R} \to [0,1]$ with α simply by taking

$$M(\xi) = \alpha(\{\xi\}) \qquad \text{for every } \xi \in \mathbb{R}.$$

Find a measure $\mu : \wp(\mathbb{R}) \to [0, \infty)$ such that

$$\alpha(A) = \int_A M \qquad \text{for every } A \subset \mathbb{R}$$

where the integral on the right is an integral with respect to (your) measure μ.

Problem 7 (probability density function, statistical values) Consider the mass density function (MDF) $\delta : \mathbb{R} \to [0, \infty)$ by

$$\delta(\omega) = 2\omega \ \chi_{[0,1]}(\omega).$$

(a) The measure π:

- **(i)** Write down a formula for $\pi([a, b])$ where $\pi : \mathcal{M} \to [0, 1]$ is the integral measure with MDF δ and $0 \leq a < b \leq 1$. (Write your answer in terms of a Riemann integral.)
- **(ii)** Show/verify that π is a probability measure.
- **(iii)** Illustrate the measure value $\pi([0.3, 1.5])$ on the graph of δ.
- **(iv)** Plot the CMF/CDF of π.

(b) statistical values:

- **(i)** Find the mean associated with the measure π.
- **(ii)** Find the statistical range associated with the measure π.
- **(iii)** Find a normalized/balanced/translated measure $\nu : \mathcal{M} \to [0, 1]$ with mean $\mu = 0$.
- **(iv)** Find the statistical range associated with the measure ν.
- **(v)** Find the variance σ^2 associated with the measure ν.

Problem 8 (The normal distribution)

(a) What does the "D" stand for in the acronym PDF?

(b) What does the "D" stand for in the acronym CDF?

(c) The values of the MDF/PDF of the normal/Gaussian measure are numerically approximated by the function `dnorm` in the statistics package R. Use R to find approximations for the following values

- **(i)** $1/\sqrt{\pi}$
- **(ii)** $e^{-2}/\sqrt{2\pi}$.
- **(iii)** $1/\sqrt{2\pi e}$.

(d) The values of the CMF/CDF of the normal/Gaussian measure are numerically approximated by the function `pnorm` in the statistics package R.

Assume the mean height of a particular population is $\mu = 1.52$ meters and the standard deviation is $\sigma = 0.09$ meters.

Use R to find approximations for the following normal probabilities:

- **(i)** The probability that a given individual has height between 1.55 meters and 1.6 meters.
- **(ii)** The probability that a given individual has height greater than 1.55 meters.
- **(iii)** The probability that a given individual has height between 1.49 meters and 1.55 meters.

Problem 9 (exponential distribution) For this problem use the functions `dexp`, `pexp`, and `rexp` in the statistics package R.

One section of the San Andreas fault produces a magnitude 6.0 or greater earthquake on average once every 20 years. Taking $\mu = 20$ as the mean in the exponential probability measure, complete the following:

(a) The probability the waiting time between (these) earthquakes is 20 years to the nearest year is approximated by the quantity

$$2\,\texttt{dexp(20, rate = }\lambda\texttt{)}.$$

(i) What value of λ would you use in this calculation?

(ii) What is the value of the approximation?

(iii) How good is the approximation? Hint: What does the (first) "`p`" in `pexp` stand for?

(b) Compile a table of values of the MDF/PDF corresponding to times

$$t = 5, 10, 15, 20, 25, 30.$$

(c) Simulate 12 wait times (between earthquakes) using `rexp`.

(d) What does the "`r`" in `rexp` stand for?

Problem 10 (memory loss) Recall that the MDF/PDF for the exponential probability measure is given by

$$\delta(\omega) = \lambda\ e^{-\lambda\omega}.$$

Let γ denote the exponential probability measure.

(a) Consider $I = [T, \infty)$.

> **(i)** Write the value $\gamma(I)$ as an improper Riemann integral.
>
> **(ii)** Evaluate the integral from part **(i)** to find $\gamma(I)$.
>
> **(iii)** Interpret the number $\gamma(I)$ as a probability.

(b) Define the **probability restriction measure** ρ_I determined by γ where I is the interval given in part **(a)**.

(c) Find a formula for the CMF/CDF of γ.

(d) Find a formula for the CMF/CDF of ρ_I where ρ_I is the probability restriction measure from part **(b)**.

(e) Show that for $t > 0$,

$$\rho_I([T + t, \infty)) = \gamma([t, \infty)) \tag{4}$$

where ρ_I is the probability restriction measure from part **(b)**.

(f) Condition (4) is said to express the fact that γ is **memoryless**. Explain in words what this means.