Worksheet 0: Some Probability

Name: Due: August 31, 2022

For the following questions, you will be working with probability measure $\mathbb{P}_{\mathcal{X}\times\mathcal{Y}}$ on $\mathcal{X}\times\mathcal{Y} = \mathbb{R}\times\mathbb{R}$ defined by density function $f_{\mathcal{X}\times\mathcal{Y}}$:

$$f(x,y) := x \cdot y^{-1/2} \mathbb{1}_{x \geq 0, x^2+y^2 \leq 1},$$

where $\mathbb{1}_{\mathrm{cond}(t)} := \begin{cases} 1 & \text{if } \mathrm{cond}(t) \\ 0 & \text{else} \end{cases}$ denotes the *indicator* function for condition $\mathrm{cond}(t)$.* Please try to fit all answers on this page.

1. Does $f_{\mathcal{X}\times\mathcal{Y}}$ define a valid pdf? If so, articulate why, else explain why not, and if not, modify f to make it a pdf. Also sketch (shading the appropriate region in $\mathbb{R}^2$) the *support* of f, i.e. the points (x,y) for which $f(x,y) \neq 0$.

2. Define the marginal density $f_{\mathcal{Y}}(y)$, express as a function of y, sketch a plot, and compute the marginal probability $\mathbb{P}_{\mathcal{Y}}(y = 0)$.

3. Components x, y define random variables $\mathbb{R}^2 \xrightarrow{\pi_\cdot} \mathbb{R}$ by projection, $\pi_x : (x,y) \mapsto x$ and $\pi_y(x,y) \mapsto y$. Are these random variables independent? Justify your answer.

4. Consider data set $S = \{(x_1,y_1),\ldots,(x_m,y_m)\} \subset \mathcal{X}\times\mathcal{Y}$. Define the covariance $\mathrm{cov}(\pi_x,\pi_y)$ and write an expression approximating this quantity from S. Each term in this expression must be computable from the data!

5. Suppose you conduct an experiment under supposition that data is $\mathbb{P}_{\mathcal{X}\times\mathcal{Y}}$-distributed, and you record a value $x = 1$ (you forgot to record the corresponding value of y). Report the p-value.

6. Consider, now, probability space $\mathcal{X} := \prod_{j=1}^{n} \mathcal{X}_j$ with joint measure $\mathbb{P}_{\mathcal{X}}$, and suppose that for any subset (projection) $\{i_1,\ldots,i_k\} \subsetneq \{1,\ldots,n\}$ setting $\mathcal{X}' := \prod_{j=1}^{k} \mathcal{X}_{i_j}$, we have equality of marginals $\mathbb{P}_{\mathcal{X}'} = \prod_{j=1}^{k} \mathbb{P}_{\mathcal{X}_{i_j}}$. Are projections $\pi_1,\ldots,\pi_k$ then independent? Why or why not? (Rationale for your answer may take the form of proof, example/counterexample, "just the definition," picture, or something else.)

*While $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, we distinguish components with notation to e.g. distinguish the marginals $f_{\mathcal{X}}$ and $f_{\mathcal{Y}}$.