Joshua Krachman
Biomedical Data Science Quiz 5

Question 1:

$Work\ for\ A$:

$15{,}932\ patients\ w/\ sepsis$
$3{,}507\ developed\ septic\ shock$

$$Adult\ prevalance = \frac{3507}{15932} = 22.01\%$$

$$Pediatric\ Prevalance = \frac{22.01\%}{3} = 7.34\%$$

$TP = 0.88 * 7.34 = 6.46\%$
$FP = 0.16 * (1 - 7.34\%) = 14.83\%$
$FN = 0.12 * 7.34 = 0.88\%$
$TN = 0.84 * (1 - 7.34\%) = 77.83\%$

	Positive Reality	Negative Reality
Positive Guess	$0.88 * 7.34 = 6.46\%$	$0.16 * (1 - 7.34\%) = 14.83\%$
Negative Guess	$FN = 0.12 * 7.34 = 0.88\%$	$TN = 0.84 * (1 - 7.34\%) = 77.83\%$

$$PPV = \frac{Sens * Prev}{(Sens * Prev) - (1 - Spec)(1 - Prev)}$$

$$PPV = \frac{TP}{TP + FP} = \frac{6.46\%}{6.46\% + 14.83\%}$$

ANSWER: Average PPV for Pediatrics = 30.34%

$Work\ for\ B$:

$$NPV = \frac{TrueNegative}{TrueNegative + FalseNegative} = \frac{77.83\%}{77.83\% + 0.88\%}$$

ANSWER: $Average\ NPV\ for\ Pediatrics = 98.88\%$

$Work\ for\ C\ on\ Next\ Page$:

$Discover\ how\ the\ features\ used\ in\ the\ Liu\ algorithm\ might\ need\ to\ be\ modified\ if\ the\ algorithm\ is\ to\ be\ applied\ to\ children.$

Work for C : According to Lecture 3 PowerPoint:

1. GCS
2. Lactate
3. BUN
4. Heart Rate
5. SOFA (CV)
6. SOFA (Respiratory)

Advantages of This Approach

Interpretable Decision Rule !

	Variable	Coefficient	SE
	Intercept	-2.974330922	0.009551231
Numeric	HR	0.292503877	0.006985994
	SBP	-0.252831939	0.006893423
	PaO_2	0.429659115	0.00438058
EHR	GCS	-0.634782631	0.005989372
	Lactate	0.733257393	0.00661599
	BUN	0.225824586	0.006995195
	WBC	0.285014711	0.005484431
	SOFA (Respiratory)	0.231603804	0.005592707
	SOFA (Coagulatory)	0.284541533	0.006787376
	SOFA (CV)	0.614226176	0.005742869

1. As https://www.brainline.org/article/what-glasgow-coma-scale states, "The GCS is usually not used with children, especially those too young to have reliable language skills."

- Therefore, the GCS score cannot be a regressor in the model with an associated weight. Instead, the PGCS, which is the pediatric version, could be used. It uses the same three visual, verbal, and motor responses, but has a different scale for scoring. Therefore, we should eliminate GCS as a regressor (set the weight to 0) and include PGCS.

2. In the paper "Arterial Blood Gases and Acid-Base Balance in Normal Children," by Donald E. Cassels and Minerva Morse, this chart shows that adults have much more consistent lactate levels than children. Therefore, using lactate as a regressor in children may be much harder to do, because the values can fluctuate so much more and still be considered normal. Therefore, we should rerun the network to lower the weight of the regressor or also normalize the lactate level per patient over time as opposed to across patients.

TABLE II

*Serum electrolyte concentrations in the arterial blood of normal boys 10 to 17 years of age**

Age	No. of cases	Bicarbonate	Chloride	Proteinate	$HCO_3^- + Cl^- + Prot.^-$	Phosphate	Lactate	Sodium
yr.		*mEq./L.*	*mEq./L.*	*mEq./L.*	*mEq./L.*	*mEq./L.*	*mEq./L.*	*mEq./L.*
10–12	11	23.0 (20.7–24.9)	104.2 (99.4–107.6)	16.7 (14.6–19.4)	143.9 (141.3–146.0)	2.55 (10) (2.25–3.05)	1.5 (0.6–2.1)	141.1 (8) (136.4–143.8)
13–14	20	23.7 (21.6–26.0)	104.3 (101.4–107.6)	16.7 (14.9–18.3)	144.7 (142.7–146.8)	2.61 (12) (2.10–3.05)	1.6 (0.9–2.8)	140.3 (12) (136.2–143.4)
15–17	33	24.3 (22.2–27.0)	104.3 (100.0–107.5)	17.2 (15.4–19.0)	145.9 (142.4–149.9)	2.45 (29) (1.91–3.07)	1.5 (0.9–2.8)	141.0 (24) (136.6–145.5)
Adult†	12	25.1 (23.9–26.0)	104.6 (103.7–105.7)	17.4 (15.9–18.3)	147.1 (144.8–149.6)		1.4 (1.2–1.9)	140.0 (131.9–143.1)

* The upper value in each case represents the mean, and quantities below in parentheses, the range of variation. In the case of phosphate and sodium, numbers in parentheses to the right of the mean indicate the number of cases included in the mean.

† Data taken from a study by Dill, Edwards and Consolazio (19).

3. BUN or Blood Urea Nitrogen test

- As https://www.mayoclinic.org/tests-procedures/blood-urea-nitrogen/about/pac-20384821, states "Urea nitrogen levels tend to increase with age." Clearly, the coefficient for this regressor may need to be altered to account for the fact that depending on the age of the child, they may have nitrogen levels that may be considered normal for an adult but actually very high for their age.

4. Heart Rate is increased in babies so the weights for heart rate would have to change to reflect this. Weights would increase because we can reliably measure the heart rates, while other regressors we cannot.

5. SOFA (CV) Neonates cannot as easily increase stroke volume during periods of high stress, thus it would make the weight for SOFA (CV) lower and increase weight for things that can reliably be measured like heart rate.

6. SOFA (Respiratory) Neonates can more easily have airways blocked because of smaller body parts, including head and trachea. This lowers the weight for this regressor, as it may not be as indicative of sepsis, as they are already more variable.

Question 2:

$$Work:$$

$$Each\ minute: 60\ 0's\ and\ 1's (0 = did\ not\ occr)$$

$$0 = Arythmia\ Event\ did\ not\ occur\ during\ each\ interval\ of\ one\ second\ duration$$

$$1 = Arythmia\ Event\ did\ occur\ during\ each\ interval\ of\ one\ second\ duration$$

$$Derive\ maximum\ likelihood\ estimate\ of\ arrythmia\ rate\ for\ each\ minute.$$

$$P(N = n) = \frac{(\lambda \Delta t)^n}{n!} e^{-\lambda}, if\ delta\ t = 1 \ldots$$

$$P(N = n) = \frac{(\lambda)^n}{n!} e^{-\lambda}$$

$$Likelihood: L(\lambda) = \prod_{n=1}^{k} P(N = n)$$

$$L(\lambda) = \prod_{n=1}^{k} \frac{(\lambda)^n}{n!} e^{-\lambda}$$

Because L is monotonically increasing, take max of log.

$$l(\lambda) = \prod_{n=1}^{k} \ln\left(\frac{(\lambda)^n}{n!} e^{-\lambda}\right)$$

$$l(\lambda) = \sum_{i=1}^{k} n ln(\lambda) - \lambda ln(e) - ln(n!)$$

$$l(\lambda) = \sum_{i=1}^{k} n ln(\lambda) - \lambda - ln(n!)$$

$$l(\lambda) = \sum_{i=1}^{k} n ln(\lambda) - \sum_{i=1}^{k} ln(n!) - k\lambda$$

$$\frac{dl}{d\lambda} = \frac{1}{\lambda} \sum_{i=1}^{k} n - k = 0\ to\ find\ max$$

$$\frac{1}{\lambda} \sum_{i=1}^{k} n = k$$

$$\lambda^* = \frac{1}{k} \sum_{i=1}^{k} n$$

$$\lambda^* = \bar{n}$$

MLE of λ = mean of n.