NAME: Galen Byrd
DATE: 1/27/19
Homework: 02
STAT/CS 387

Problem 1

In the context of decision analysis, “the map is not the territory” means that the models that we build are helpful in guiding us towards a result, but do not span the whole problem space. In other words, the “map” or “model” that we create is a scaled-down representation of reality and does not contain all the complexities of the whole problem. Just as a map cannot tell you the exact slope of the landscape at every point, but it can help approximate the slope of a hill in general. A map can also be useful in finding a general route from point A to B, but it will not be specific enough to indicate the exact amount of inches to walk in each direction.

While decision analysis might be one of the most rational ways to make a decision, there are still many ways it could fail. One obvious way a decision analysis could fail is if we are given inaccurate data. If we build a model based on bad data then every decision made using that model will now be mistaken by the error in the model on top of the error in the data. This is true with all models, but in decision analysis could be more costly depending on the risk associated with the decision. In addition to bad data, if we build a model without all the variables necessary to generalize a decision, the analysis will fail. For example, if we are trying to decide which horse to bet on in a race and our model tells us that horse A will almost always beat horse B then we would place a bet on horse A. What if weather was not included in the model and horse B performs much better than horse A in the rain. This would mean that our decision analysis would tell us to bet for horse A when in this case horse B would have a higher utility given a more representative model. Another example of this would be using google maps to plan a driving route. If the route finder does not include the traffic at the current time of day, then it may send you on a route that takes longer than the optimal route at that time. The more complex our model is, the more precise, but less generalizable it will become.

Another major way a decision analysis could fail is if a potential outcome is omitted. For example, in the cancer treatment example we looked at in class, one could choose Radiotherapy, Surgery or None for their treatment, but what if there was another option, say a Pill, that was not included in the analysis. As an extension of this example, consider a question with infinite outcomes. How would we create a solution space for the question, What should I do over spring break? It is impossible, as there is an infinite number of options with vastly different utilities between individuals.

Problem 2

To decide wether or not the man should choose to have a biopsy we must calculate the six utilities of R/S/N given that the man took a biopsy. We will compare the utilities of each given a positive vs negative result from the biopsy. We compute the utilities given a biopsy using Pr(cancer|B+), Pr(no cancer|B+), Pr(cancer|B-) and Pr(no cancer|B-) instead of simply Pr(cancer) and Pr(no cancer). Using probability rules Pr(cancer|B+) = 1-Pr(no cancer|B+). We uses Bayes' Theorem to find Pr(cancer|B=b) = (Pr(B=b|cancer)*Pr(cancer))/Pr(B=b) and see that Pr(B=b) = Pr(B=b|cancer)*Pr(cancer) + Pr(B=b|no cancer)*Pr(no cancer), where b are the possible results of the biopsy, positive or negative.

U(R|B+) = .997 * 16.7 + .003 * 34.8 - 1 = 15.75

U(S|B+) = 0.35 * 0 + 0.65 * (.997 * 20.3 + .003 * 34.8 - 1) = 12.57

U(N|B+) = .997 * 5.6 + .003 * 34.8 = 5.69

U(R|B-) = .734 * 16.7 + .266 * 34.8 - 1 = 20.51

U(S|B-) = 0.35 * 0 + 0.65 * (.734 * 20.3 + .266 * 34.8 - 1) = 15.05

U(N|B-) = .734 * 5.6 + .266 * 34.8 = 13.37

Given these six QALEs the man should not get the biopsy, as doing so does not change his decision given the outcome of the test. That is, given either a positive or negative test result the maximum utilities are given by getting Radiotherapy, so the biopsy does not add any information to the decision and the 5% risk of immediate death is not worth it.

Problem 3

Absolute error loss:

$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}|$, or $\theta - \hat{\theta}$ when $\theta > \hat{\theta}$, $\hat{\theta} - \theta$ when $\theta < \hat{\theta}$

$r(\hat{\theta}|x) = \int_{-\infty}^{\hat{\theta}} (\theta - \hat{\theta})\, f(\theta|x)\, d\theta + \int_{\hat{\theta}}^{\infty} (\hat{\theta} - \theta)\, f(\theta|x)\, d\theta$, minimize $r(\hat{\theta}|x)$ by differentiating with respect to $\hat{\theta}$ and setting equal to zero. Using Leibniz rule we see:

$dr/d\hat{\theta} = \int_{-\infty}^{\hat{\theta}} f(\theta|x)\, d\theta + (\hat{\theta}-\hat{\theta})*1 - (\hat{\theta}- -\infty)*0 - \int_{\hat{\theta}}^{\infty} f(\theta|x)\, d\theta - (\hat{\theta}-\hat{\theta})*1 + (\infty-\hat{\theta})*0$

$0 = \int_{-\infty}^{\hat{\theta}} f(\theta|x)\, d\theta - \int_{\hat{\theta}}^{\infty} f(\theta|x)\, d\theta$

$\int_{-\infty}^{\hat{\theta}} f(\theta|x)\, d\theta = \int_{\hat{\theta}}^{\infty} f(\theta|x)\, d\theta$, Therefore we know $\hat{\theta}$ must split the pdf of our prior distribution $f(\theta|x)$ in half, which by definition would be the median of the distribution.

Zero-one loss:

$L(\theta, \hat{\theta}) = 0$ when $\theta = \hat{\theta}$, 1 when $\theta \neq \hat{\theta}$

Considering a discrete case we see $E[L(\theta, \hat{\theta})|X] = \sum_{\theta \in \boldsymbol{\theta}} (\hat{\theta} \neq \theta) * f(\theta|x)$

$= 1 - \sum_{\theta \in \boldsymbol{\theta}} (\hat{\theta} = \theta) * f(\theta|x)$, which simplifies to $f(\hat{\theta}|x)$ as $f(\theta|x)$ is only added when $\theta = \hat{\theta}$

$= 1 - f(\hat{\theta}|x)$

We then wish to maximize the posterior probability $f(\hat{\theta}|x)$, which happens when $\hat{\theta}$ is the mode of the posterior.