Decision Theory Assignment 1

Omkar Bhutra (omkbh878)

1 January 2020

Assignment 1:

1a

Using Likihood as a measure of how likely is an event is a matter of inference to the best explanation. Inference to the best explanation:

$P(B|\overline{A}, I) = 0.1 \ P(B|A_1, I) = 20 \ P(B|A_2, I) = 10 \ P(B|A_3, I) = 2$

$P(B|A_1, I) > P(B|A_j, I)$ Then A_1 is considered best explanation for B and is provisionally accepted Hence, GP's diagnosis using the principles of inference to the best explanation is that disease A_1 explains the Symptoms.

1b-i

$P(A_1) = 0.2 \ P(A_2) = 0.3 \ P(A_3) = 1 \ P(B|\overline{A}) = 0.1$

Conditional probabilities: $P(A_k|B) = \dfrac{P(B|A_k).P(A_k)}{P(B|A_k).P(A_k)+P(B|A_j).P(A_j)+P(B|A_i).P(A_i)+P(B|\overline{A})(1-[P(A_k)+P(A_j)+P(A_i)])}$

$P(A_1|B) = \dfrac{0.2 X 0.002}{(0.2 X 0.002)+(0.1 X 0.003)+(0.02 X 0.01)+(0.001)(1-[0.002+0.003+0.01])}$ = 21.220%

$P(A_2|B) = \dfrac{0.1 X 0.003}{(0.2 X 0.002)+(0.1 X 0.003)+(0.02 X 0.01)+(0.001)(1-[0.002+0.003+0.01])}$ = 15.915%

$P(A_3|B) = \dfrac{0.02 X 0.01}{(0.2 X 0.002)+(0.1 X 0.003)+(0.02 X 0.01)+(0.001)(1-[0.002+0.003+0.01])}$ = 10.610%

$P(\overline{A}|B) = \dfrac{P(B|\overline{A}).P(\overline{A})}{P(B|A_k).P(A_k) + P(B|A_j).P(A_j) + P(B|A_i).P(A_i) + P(B|\overline{A})(P(\overline{A}))} = \dfrac{0.001 X (1 - [0.002 + 0.003 + 0.01])}{(0.2 X 0.002) + (0.1 X 0.003) + (0.02 X 0.01) + (0.001)(1 - [0.$

= 52.631%

1b-ii

Since, $P(\overline{A}|B) > P(A_1|B) > P(A_2|B) > P(A_3|B)$ Then $\overline{A}$ is considered best explanation (Having no disease) for the symptoms and is provisionally accepted.

2-i

Posterior probability model for Binomial sampling: The sample size n could be fixed, in which case $\tilde{r}$ is a random variable and the distribution of $\tilde{r}$ given p and n is a binomial distribution.

$$P(p|n,r) \propto \begin{bmatrix} n \\ r \end{bmatrix} p^r (1-p)^{n-r} P(p)$$

For Pascal sampling, the number of successes r could be fixed (sample until r successes are obtained), in which case $\tilde{n}$ is a random variable and the distribution of $\tilde{n}$ given p and r is a Pascal distribution, which is of the following form:

$$P(p|n,r) \propto \binom{n-1}{r-1} p^r (1-p)^{n-r}.P(p)$$

These two distributions differ only in the first terms, the combinatorial terms. The reason for this is that the last trial must be a success in Pascal sampling. In the posterior, The combinatorial terms are irrelevant since they do not involve $\tilde{p}$. The combinatorial term can be moved outside the integral where it appears in both the numerator and denominator and can be cancelled out. Therefore, the liklihood function fot the Bernoulli process can be taken equal to $p^r (1-p)^{n-r}$ so that it isnt necessary to know whether the sampling is done with n fixed or with r fixed.

$$f''(p|y) = \frac{f'(p)f(y|p)}{\int_0^1 f'(p)f(y|p)dp} = \frac{f'(p)P(r|n,p)}{\int_0^1 f'(p)P(r|n,p)dp}$$

$$= \frac{\frac{(n'-1)!}{(r'-1)!(n'-r'-1)!}p^{r'-1}(1-p)^{n'-r'-1}\binom{n}{r}p^r(1-p^{n-r})}{\int_0^1 \frac{(n'-1)!}{(r'-1)!(n'-r'-1)!}p^{r'-1}(1-p)^{n'-r'-1}\binom{n}{r}p^r(1-p^{n-r})dp}$$

The terms not involving p can be moved outside the integral

$$f''(p|y) = \frac{p^{r'-1}(1-p)^{n'-r'-1}p^r(1-p)^{n-r}}{\int_0^1 p^{r'-1}(1-p)^{n'-r'-1}p^r(1-p)^{n-r}dp}$$

$$f''(p|y) = \frac{1}{k}p^{r''-1}(1-p)^{n''-r''-1}$$

$$k = \frac{(r''-1)!(n''-r''-1)!}{(n''-1)!}$$

Either case, the term cancels out so the final result is the same

2-ii

To determine the posterior distribution, the only information needed about the sample consists of the values of the sample statistics, r and n. Hence, r and n are sufficient statistics. The procedure is used to tell statisticians when to stop sampling is called the stopping rule, if the stopping rule has no effect on the posterior distribution, then it is said to be non-informative. The two stopping rules are: (sample until n trails and sample

until r successes) are both non-informative.

3

Prior probabilities:

$$P(\lambda = 30) = P(\lambda = 50) = 0.25,$$
$$P(\lambda = 50) = 0.50$$

Liklihoods:

$$L(\lambda = 30, r = 380, t = 2) = \frac{(\lambda.t.5)^r.e^{-\lambda.t.5}}{r!} = \frac{(30.2.5)^{380}e^{-30.2.5}}{380!} = 1.103X10^{-6}$$

$$L(\lambda = 50, r = 380, t = 2) = \frac{(50.2.5)^{380}e^{-50.2.5}}{380!} = 3.0645X10^{-9}$$

Therefore,

$$L(\lambda = 40, r = 380, t = 2) = \frac{(40.2.5)^{380}e^{-40.2.5}}{380!} = 1.23X10^{-2}$$

Posterior probabilties:

$$P(\lambda|r = 380, t = 2) = \frac{\frac{(2\lambda.5)^{380}e^{-2\lambda.5}}{380!}.P(\lambda)}{\sum(Liklihoods X P(\lambda))}$$

The λ^r terms cancel out

$$P(\lambda = 30|r = 380, t = 2) = \frac{((30.2.5)^{380}e^{-30.2.5})\frac{1}{4}}{((30.2.5)^{380}e^{-30.2.5}.\frac{1}{4}) + ((40.2.5)^{380}e^{-40.2.5}.\frac{1}{2}) + ((50.2.5)^{380}e^{-50.2.5}.\frac{1}{4})}$$
$$= 4.475X10^{-5}$$

$$P(\lambda = 40|r = 380, t = 2) = \frac{((40.2.5)^{380}e^{-40.2.5})\frac{1}{2}}{((30.2.5)^{380}e^{-30.2.5}.\frac{1}{4}) + ((40.2.5)^{380}e^{-40.2.5}.\frac{1}{2}) + ((50.2.5)^{380}e^{-50.2.5}.\frac{1}{4})}$$
$$= 0.999$$

$$P(\lambda = 50|r = 380, t = 2) = \frac{((50.2.5)^{380}e^{-50.2.5})\frac{1}{4}}{((30.2.5)^{380}e^{-30.2.5}.\frac{1}{4}) + ((40.2.5)^{380}e^{-40.2.5}.\frac{1}{2}) + ((50.2.5)^{380}e^{-50.2.5}.\frac{1}{4})}$$
$$= 1.2458X10^{-7}$$

4-i

A family of probability distributions belongs to the k-parameter exponential class of distributions if the probability density can be written as:

$$f(x|\theta) = e^{\sum_{j=1}^{k} A_j(\theta)B_j(x)+C(x)+D(\theta)}$$

where:

$$\theta = (\theta_1, \ldots, \theta_k)$$

$$A_1(\theta), \ldots, A_k(\theta)$$

and $D(\theta)$ are functions of the parameter θ only (and not of x) $B_1(\theta), \ldots, B_k(\theta)$ and $C(x)$ are functions of x only (and not of θ)

Two parameter Beta distribution , shape parameterisation:

$$f(x|\theta) = f(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

$$f(x|\theta) = f(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}\gamma(\alpha+\beta)}{\gamma(\alpha)\gamma(\beta)}$$

where:

$$B(\alpha, \beta) = \frac{\gamma(\alpha)\gamma(\beta)}{\gamma(\alpha+\beta)}$$

$$= e^{(\alpha-1)(\ln(x))+(\beta-1)\ln(1-x)+\ln(\gamma(\alpha+\beta))-\ln(\gamma(\alpha)-\ln(\gamma(\beta))}$$

Parametric form:

$$A_1(\theta) = \alpha - 1$$

$$B_1(x) = \ln(x)$$

$$A_2(\theta) = \beta - 1$$

$$B_2(x) = \ln(1-x)$$

$$C(x) = 0$$

$$D(\theta) = \ln(\gamma(\alpha+\beta)) - \ln(\gamma(\alpha) - \ln(\gamma(\beta))$$

When the sample distribution is Binomial such as $X \sim Bin(n, \pi)$ then, The Conjugate Prior is a Beta distribution $\pi \sim \beta(\alpha, \beta)$ and the posterior is also a Beta distribution such as $\pi|x \sim Beta(\alpha + x, \beta + n - x)$

pdf of sample point distribution:

$$f(x|\theta) = e^{\sum_{j=1}^{k} A_j(\theta)B_j(x)+C(x)+D(\theta)}$$

liklihood:

$$\prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} e^{A_j(\theta)B_j(x_i)+C(x_i)+D(\theta)}$$

$$= e^{\sum_{i=1}^{n}(\sum_{j=1}^{k} A_j(\theta)B_j(x_i)+C(x_i)+D(\theta))}$$

$$= e^{\sum_{j=1}^{k} A_j(\theta) \sum_{i=1}^{n} B_j(x_i)+\sum_{i=1}^{n} C(x_i)+n.D(\theta)}$$

prior density:

$$p(\theta|\alpha_1, \ldots, \alpha_k, \alpha_{k+1})$$

$$= e^{\sum_{j=1}^{k} A_j(\theta).\alpha_j+\alpha_{k+1}.D(\theta)+K(\alpha_1,\ldots,\alpha_k,\alpha_{k+1})}$$

$$\propto e^{\sum_{j=1}^{k} A_j(\theta).\alpha_j+\alpha_{k+1}.D(\theta)}$$

4-ii

posterior distribution:

$$q(\theta|x, \alpha_1, \ldots, \alpha_k, \alpha_{k+1})$$

$$= q(\theta|x_1, \ldots, x_n;, \alpha_1, \ldots, \alpha_k, \alpha_{k+1})$$

$$\propto \prod_{i=1}^{n} f(x_i|\theta). \, p(\theta|\alpha_1, \ldots, \alpha_k, \alpha_{k+1})$$

$$= e^{\sum_{j=1}^{k} A_j(\theta) \sum_{i=1}^{n} B_j(x_i)+\sum_{i=1}^{n} C(x_i)+n.D(\theta)} . e^{\sum_{j=1}^{k} A_j(\theta).\alpha_j+\alpha_{k+1}.D(\theta)+K(\alpha_1,..,\alpha_k,\alpha_{k+1})}$$

$$\propto e^{\sum_{j=1}^{k} A_j(\theta)(\sum_{i=1}^{n} B_j(x_i)+\alpha_j)+(n+\alpha_{k+1})D(\theta)}$$

posterior distribution hyperparameters:

$$\alpha_1 + \sum_{i=1}^{n} B_1(x_i), \ldots, \alpha_k + \sum_{i=1}^{n} B_k(x_i), \alpha_{k+1} + n$$