CS 434 HW0

Tristan Gavin

March 31, 2021

Probability

1. Bayes Theorem and Marginalization

Information we are given:

- rains 73 days each year (365 days)
- Weatherperson is 70% accurate forcasting rain when it rains
- Weatherperson innacurately predicts rain 30% of the time when it does not rain

Lets create the set $R \in \{0, 1\}$ for when it doesn't rain and when it does rain. The set $F \in \{0, 1\}$ for whether the weatherperson doesn't predict rain or does predict rain respectively. We want to find the probability that it will rain given that the weatherperson predicted rain (denoted $P(R = 1|F = 1)$). To solve this we will implement Bayes Theorem

$$P(R=1|F=1) = \frac{P(F=1|R=1)P(R=1)}{P(F=1|R=1)P(R=1) + P(F=1|R=0)P(R=0)}$$

Fortunately we are given enough information to fill out this equation.

$P(F = 1|R = 1) = 0.7$ (probability that the forcaster predicted rain given that it rains)

$P(R = 1) = 73 \div 365 = 0.2$ (probability of rain on any given day)

$P(F = 1|R = 0) = 0.3$ (Probability that weatherperson innacurately predicts rain given it does not rain)

$P(R = 0) = 1 - (73 \div 365) = 0.8$ (probability that it doesn't rain on any given day)

$$P(R=1|F=1) = \frac{(0.7)(0.2)}{(0.7)(0.2) + (0.3)(0.8)} = \frac{0.14}{0.38} \approx 0.3684$$

There is a 36.84% chance that it will rain tomorrow

2. Computing Expected Values from Discrete Ditributions

There is a $P(x = 1) = \frac{1}{6}$ chance of rolling a one and $p(x \neq 1) = \frac{5}{6}$ chance of not rolling a 1. Taking the expected values of all sample events and multiplying them by the amount of money we will make or lose if the event is to happen we have $P(x = 1)(1) + P(x = 2)(-.25) + P(x = 3)(-.25) + P(x = 4)(-.25) + P(x = 5)(-.25) + P(x = 6)(-.25) = 1/6(1.25) + 1/6(-.25) + 1/6(-.25) + 1/6(-.25) + 1/6(-.25) + 1/6(-.25) = -.25$ we will expect to lose on average 25 cents each time so no this is not a good bet.

3. Linearity of Expectation

For this problem we will solve the integral using these properties of variance and probability density functions.

$\int_{-\inf}^{\inf} p(x)dx = 1$ (1)

$\mu = \int_{-\inf}^{\inf} xp(x)dx$ (2)

$\sigma^2 = \int_{-\inf}^{\inf} x^2 p(x)dx - \mu^2$ (3)

First we will distribute the p(x) across the polynomial to get

$$\int_{-\inf}^{\inf} ax^2p(x) + bxp(x) + cp(x)dx = a\int_{-\inf}^{\inf} x^2p(x) + b\int_{-\inf}^{\inf} xp(x) + c\int_{-\inf}^{\inf} p(x)$$

The answer falls out with the properties written above. The first term is $\sigma^2 a = (1)a = a$ (3). Since we are given in the problem that $\mu = 0$ the second term is $\mu = (0)b = 0$ (2). And the third term is just the integral of a pdf (1) times c which is c. As our final answer we have

$$\int_{-\inf}^{\inf} p(x)(ax^2 + bx + c)dx = a + c$$

4. Cumulative Density Functions / Calculus

We need to use the cumulative density function and integration to find the area under the curve for when $0 \leq x \leq \frac{1}{2}$ and $\frac{1}{2} \leq x \leq 1$ and add the two together to get the total area under the curve.

$$\int_0^{1/2} 4xdx + \int_{1/2}^{1} (-4x + 4)dx = 1$$

To find $P(a < X < b)$ we just put a and b into the integrals respectively depending on if their value is less than or greater than 1/2. (I think this is all we need to show)

Linear Algebra

1. Transpose and Associative Property

For this proof we will use properties of transpose and some basic linear algebra manipulations to get two identical scalars mutliplied by eachother.

$$\begin{aligned} x^TBx &= x^Tbb^Tx && \text{(definition of B given in problem)} \\ &= (b^Tx)^T(b^Tx) && \text{(definition matrix transpose)} \\ &= [(x^Tb)(x^Tb)]^T && \text{(definition matrix transpose)} \end{aligned}$$

since $x, b \in R^{d\times 1}$ then (x^Tb) has dimensions $(1 \times d) \times (d \times 1)$ which by definition of matrix multiplication has a dimension of (1×1) in otherwords just a scalar. Lets call this scalar c Then we have $[(c)(c)]^T = (c^2)^T = c^2$ (by definition of transpose on a scalar) $c^2 = x^TBx \geq 0$ as was to be shown.

2. Solving System of Linear Equation with Matrix Inverse

(a)

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix}$$

(b)

Using software we find the inverse of A to be: $\begin{bmatrix} -1 & .5 & .5 \\ 1 & -.5 & 0 \\ 2 & -.5 & -1 \end{bmatrix}$ We can now use this information to find x by solving the equation $A^{-1}b = x$:

$$\begin{bmatrix} -1 & .5 & .5 \\ 1 & -.5 & 0 \\ 2 & -.5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = x$$

Proving Things

1. Finding Maxima of a Function

For this proof we will consider the graph $f(x) = ln(x) - x + 1$ and show that its maximum value for $x > 0$ is $f(1) = 0$. First we take the derivitive of our function $f(x)$ and find where it is equal to 0.

$$f'(x) = \frac{1}{x} - 1 = 0 \implies x = 1 \implies f(1) = 0 \text{ is a maxima}$$

We know there is a maxima or minima at $x = 1$ to find which it is lets look at the second derivitive.

$$f''(x) = -\frac{1}{x^2}$$

Since $f''(x) < 0$ for all $x > 0$ we know that $x = 1$ must be a maximum point on the interval $(0, \infty)$. Then $ln(x) - x + 1 \leq 0 \implies ln(x) \leq x - 1$ for all $x > 0$ as was to be shown.

2. Proving Abstract Concepts

We need to show that the divergence between p_i and q_i is non-negative given that $p_i \geq 0, q_i \geq 0,$ for all $i \in \{0, 1, 2 \ldots k\}$ i.e $\sum_{i=1}^{k} p_i \ln \frac{p_i}{q_i} \geq 0$

$$\begin{aligned}
&\sum_{i=1}^{k} p_i \ln \frac{p_i}{q_i} \\
&= \sum_{i=1}^{k} p_i(-\ln \frac{q_i}{p_i}) \\
&\geq \sum_{i=1}^{k} p_i(-\frac{q_i}{p_i} + 1) \text{ (this comes from problem 3.1 innequality changes due to negative)} \\
&= -\sum_{i=1}^{k} q_i + \sum_{i=1}^{k} p_i \\
&= -1 + 1 = 0
\end{aligned}$$

Debriefing

1. I spent probably 8 hours on this assignment
2. moderate difficulty
3. I worked on all of it alone except for some help from a math tutor on a few probability questions
4. 80%
5. probably spent 6 hours working on the problems and another two hours trying to remember latex commands.