09/16

Aksel Kretsinger-Wolters

1) A = Alice gets wet
M = It rains on Monday
T = It rains on Tuesday

$p_M = 0.75$

$p(T \mid M=1) = 0.5$
$p(T \mid M=0) = 0.2$

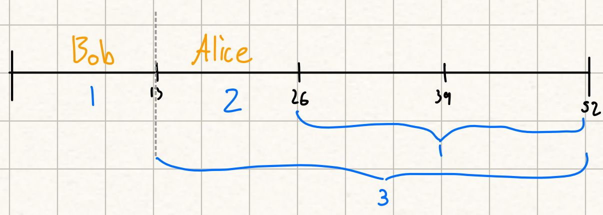
$p(A \mid M=1) = p(A \mid T=1) = 0.7$

$p(A_M \& A_T) = ?$

$$p(A_M \& A_T) = 0.75 \times 0.7 \times 0.5 \times 0.7$$
$$= \tfrac{3}{4} \times \tfrac{1}{2} \times 0.7^2$$
$$= \tfrac{3}{8} \times \tfrac{49}{100} = \tfrac{147}{800}$$

2) A B T H

a) $P(A=2 \mid B=1)$



B_1 = Bob has 1 ace

A_2 = Alice has 2 aces

$$P(A_2 \mid B_1) = \frac{\binom{3}{2}\binom{36}{11}}{\binom{39}{13}}$$

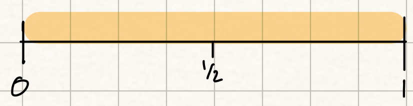
b) C = # clubs dealt to Alice
S = # spades dealt to Alice

$$E(C) = \tfrac{13}{4}$$

$$E(C \mid S) = (13 - S) \times \frac{13}{39}$$

$$= \frac{13 - S}{3}$$

3)



Essentially, this question boils down to the probability that a given offer is $> X_1$

Because the distribution is uniform continuous, we can ignore ties

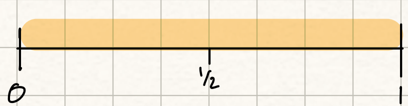
$\mathbb{E}[X_1] = 1/2$

$p[X_i > X_1] = 1/2$

$\mathbb{E}[\text{draws}] = 1/p_{\text{success}} = \frac{1}{1/2} = 2$

$\mathbb{E}[\text{offers}] = 1 + \mathbb{E}[\text{draws}] = 1 + 2 = 3$

4)



(a) $\mathbb{E}[X_1] = 1/2$

$p[X_i > X_1] = 1/2$

$E[Y_n] = \frac{n-1}{2}$

(b) $E[Z_n] = \frac{n-1}{2}$, same solution as 4a

function. What is $E[Z_n]$? Also, prove that with probability at least $1-\delta$, for any $\delta \in (0,1)$, we have that $Z_n \le \frac{n}{2} + O(\sqrt{n \log(1/\delta)})$. That is,

$$\Pr\left(Z_n \le \frac{n}{2} + O(\sqrt{n \log(1/\delta)})\right) \ge 1 - \delta$$

For the proof, we'll start by showing the variance and stdev of this process

$p_{accept} = 1/2$

$Var(p,n) = n \cdot p \cdot (1-p)$

$= n \cdot \left(\frac{1}{2}\right)^2$

$= \frac{n}{4}$

$Std(p,n) = \sqrt{Var(p,n)} = \sqrt{\frac{n}{4}} = \sqrt{n}/2$

$$\mathbf{Pr}[|M - \mathbf{E}[M]| > \alpha] \le 2\exp\left(\frac{-\alpha^2}{4\sum_{i=1}^{r} \mathbf{Var}[X_i]}\right).$$

$$P\left(\left|Z_n - \frac{n-1}{2}\right| > O\left(\sqrt{n \log(1/\delta)}\right)\right) \le 2 \times \exp\left\{\frac{-O\left(\sqrt{n\log(1/\delta)}\right)^2}{4 \times \frac{n}{4}}\right\}$$

$$\le 2\exp\left\{\frac{-\cancel{n}\log(1/\delta)}{\cancel{n}}\right\}$$

$$\le 2 \times -(1-\delta)$$

$$P\left(Z_n \le \frac{n-1}{2} + O\left(\sqrt{n\log(1/\delta)}\right)\right) \ge 2(1-\delta)$$

c) $\mathbb{E}[W_n] = \frac{n-1}{2}$

There are two seperate intuitive explanations that I will offer

1) If you collected all n data points and took the mean, you would expect ½ of those points to reside above the mean (due to the symmetry of the uniform distibution: mean = median)

The introduction of ordering does <u>not</u> impact this estimate, because the question establishes the independence of each data point

2) If you view this process one-by-one, a priori you'd expect $\mu = 1/2$

$$\mathbb{E}[U \sim (a,b)] = \frac{a+b}{2}$$

So, at any given day, your estimate for $p\left[X_n > \frac{1}{n-1}\sum_{i=1}^{n-1} x_i\right] = 1/2$

Problem 5: Linear Program Duality

Lets first group constants and variables

$$v + 0.5x_1 - 0.2x_2 - 0.3x_3 \geq 2$$
$$v - 0.1x_1 + 0.1x_2 \geq 3$$
$$v - 0.2x_1 - 0.95x_2 + 0.95x_3 \geq 5$$

Written in standard form, the coefficient vector c, coefficient matrix A, and constant vector b are:

$$c = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0.5 & -0.2 & -0.3 \\ 1 & -0.1 & 0.1 & 0 \\ 1 & -0.2 & -0.95 & 0.95 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

Dual LP Formulation

Here, the dual variable vector is $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$.

Our goal is to maximize $b^T y = 2y_1 + 3y_2 + 5y_3$.

The constraints are found by transposing A and multiplying by y, then setting it equal to c:

$$A^T = \begin{pmatrix} 1 & 1 & 1 \\ 0.5 & -0.1 & -0.2 \\ -0.2 & 0.1 & -0.95 \\ -0.3 & 0 & 0.95 \end{pmatrix}$$

The constraints are:

$$y_1 + y_2 + y_3 = 1$$
$$0.5y_1 - 0.1y_2 - 0.2y_3 = 0$$
$$-0.2y_1 + 0.1y_2 - 0.95y_3 = 0$$
$$-0.3y_1 + 0.95y_3 = 0$$

Sign Constraints: Since the primal constraints are of the "$\geq$" type, the dual variables must be non-negative: $y_1, y_2, y_3 \geq 0$.

Final Dual Formulation: The complete dual LP is:

$$\begin{aligned} \text{Maximize} \quad & 2y_1 + 3y_2 + 5y_3 \\ \text{s.t.} \quad & y_1 + y_2 + y_3 = 1 \\ & 0.5y_1 - 0.1y_2 - 0.2y_3 = 0 \\ & -0.2y_1 + 0.1y_2 - 0.95y_3 = 0 \\ & -0.3y_1 + 0.95y_3 = 0 \\ & y_1, y_2, y_3 \geq 0 \end{aligned}$$

Problem 6. (15 points) Consider the function $f(x) = \log(\sum_{i=1}^{n} e^{x_i})$.

1. What is the gradient $\nabla f(x)$ of f? What is the gradient at $x = (1, 1, \ldots, 1)$?
2. What is the Hessian $\nabla^2 f(x)$ of f? What is the Hessian at $x = (1, 1, \ldots, 1)$?

1) $f(x) = \log\left(e^{x_1} + e^{x_2} + \dots + e^{x_n}\right)$

$$\nabla f(x) = \frac{1}{\Sigma e^{x_i}} \left[e^{x_1} + e^{x_2} + \dots + e^{x_n}\right]$$

$$\nabla f(\mathbb{1}) = \frac{1}{ne} \cdot \vec{e} = \frac{1}{n}(1, 1, \dots, 1)^T$$

2) $\nabla f(x) = \frac{1}{\Sigma e^{x_i}} \left[e^{x_1} + e^{x_2} + \dots + e^{x_n}\right]$

$\nabla^2 f(x) \in \mathbb{R}^{n \times n}$

$$\frac{\partial f}{\partial x_j} = \frac{e^{x_j}}{\sum_{i=1}^{n} e^{x_i}} \qquad \frac{\partial^2 f}{\partial x_j \, \partial x_k} = \frac{\partial f}{\partial x_k}\left(\frac{e^{x_j}}{\sum_{i=1}^{n} e^{x_i}}\right)$$

<u>If $j=k$ (diagonal entries)</u>

$$\frac{\partial^2 f}{\partial x_j \, \partial x_k} = \frac{e^{x_j}\sum_{i=1}^{n} e^{x_i} - e^{x_j} \cdot e^{x_j}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} = \frac{e^{x_j}\sum_{i \neq j} e^{x_i}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} = \frac{\partial f}{\partial x_j} - \frac{e^{2x_j}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2}$$

<u>If $j \neq k$</u>

$$\frac{\partial^2 f}{\partial x_j \, \partial x_k} = \frac{0 - e^{x_j} \cdot e^{x_k}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} = \frac{-e^{x_j} \cdot e^{x_k}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} = -\frac{\partial f}{\partial x_j} \times \frac{\partial f}{\partial x_k}$$

Notice that the term getting subtracted is the "same" in both.

Lets set V_i to a common term

$$v_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

$$\nabla^2 f(x) = I \cdot v_{1\ldots n} - vv^T$$

$$J = \vec{1}\,\vec{1}^T \in \mathbb{R}^{n \times n}$$

$$\nabla^2 f(\vec{1}) = \frac{1}{n} I - \frac{1}{n^2} \cdot J$$

3. Consider the following optimization problem:

$$\begin{array}{ll} \min\limits_{x} & \log(\sum_{i=1}^{n} e^{x_i}) \\ s.t. & \sum_i x_i = 1 \end{array}$$

Argue that this is a convex optimization problem and solve it analytically using optimality conditions. State the optimal solution.

3) If we can show that the objective function we seek to minimize is convex, we can argue that the problem is convex

To prove convexity, we must normally show that the Hessian is PSD. Fortunately for us, it is known that log-sum-exp objective functions are convex

To solve, set up a LaGrangian

$$L(x, \lambda) = \log\left(\sum_{i=1}^{n} e^{x_i}\right) + \lambda\left(\sum_{i=1}^{n} x_i - 1\right)$$

$$\frac{\partial L}{\partial x_j} = \frac{e^{x_j}}{\sum_{i=1}^{n} e^{x_i}} - \lambda = 0$$

$$\frac{e^{x_j}}{\sum_{i=1}^{n} e^{x_i}} = \lambda \quad \forall j \in n$$

$$e^{x_1} = e^{x_2} = \ldots = e^{x_n}$$

$$x_1 = x_2 = \ldots = x_n$$

$$x_j = \frac{1}{n} \quad \forall j \in n$$

$$\hat{x} = \frac{1}{n} \cdot \vec{1}^T \in x^n$$