1001144647
Javier Salazar | MACHINE LEARN. | HW #1 | 1/

1. $P(S=DFW)=0.2$ $\quad P(S=Min.)=0.8$ $\quad P(T\leq 40\,|\,DFW)=0.2$
$P(T\leq 40\,|\,Min.)=0.8$

1) $P(DFW\,|\,T>40)? = \dfrac{P(S=DFW)\,P(T>40\,|\,DFW)}{P(S=DFW)\,P(T>40\,|\,DFW)+P(S=Min.)\,P(T>40\,|\,Min)}$

$P(T>40\,|\,DFW)+P(T\leq 40\,|\,DFW)=1$ $\quad\therefore\quad P(T>40\,|\,DFW)=0.8$
$P(T>40\,|\,Min)=0.2$

$\therefore = \dfrac{0.2\times 0.8}{0.2\times 0.8+0.8\times 0.2} = \boxed{\frac{1}{2}}$

2) $P(T_2>40\,|\,T_1>40) = \dfrac{P(T_2>40\cap T_1>40)}{P(T_1>40)} = \dfrac{[P(T_2>40\cap T_1>40\,|\,DFW)\,P(DFW) + P(T_2>40\cap T_1>40\,|\,Min)\,P(Min)]}{P(T_1>40)}$

$= [P(T_1>40\,|\,DFW)\cdot P(T_2>40\,|\,DFW)\cdot P(DFW) + P(T_1>40\,|\,Min)\,P(T_2>40\,|\,Min)\,P(Min)]\,/\,P(T_1>40)$

$= \dfrac{0.8\times 0.8\times 0.2+0.2\times 0.2\times 0.8}{P(T_1>40)=P(T_1>40\,|\,DFW)\,P(DFW)+P(T_1>40\,|\,Min)\,P(Min)=0.8\times 0.2+0.8\times 0.2} = \boxed{1/2}$

3) $P(T_1>40\cap T_2>40\cap T_3>40) =$

$= P(T_1>40\cap T_2>40\cap T_3>40\,|\,DFW)\,P(DFW) + P(T_1>40\cap T_2>40\cap T_3>40\,|\,Min)\,P(Min)$

$= P(T_1>40\,|\,DFW)\,P(T_2>40\,|\,DFW)\,P(T_3>40\,|\,DFW)\,P(DFW) +$
$P(T_1>40\,|\,Min)\,P(T_2>40\,|\,Min)\,P(T_3>40\,|\,Min)\,P(Min)$

$= 0.8\times 0.8\times 0.8\times 0.2 + 0.2\times 0.2\times 0.2\times 0.8 = \boxed{0.1088}$

1001144647 Javier Salazar | MACHINE LEARNING | HW #1

2. $f_{X,Y}(x,y) = K(x+y) \quad 0 \le x,y \le 1$

1) $\iint_\Omega f_{X,Y}(x,y) = 1 \therefore K\int_0^1\int_0^1 (x+y)\,dx\,dy = 1 \to K\int_0^1 \left[\frac12 x^2 + yx\right]_0^1 dy = 1$

$K\int_0^1\left[\frac12 + y\right]dy = 1 \to K\left[\frac12 y + \frac12 y^2\right]_0^1 \to K\left[\frac12+\frac12\right]=1 \therefore \boxed{K=1}$

2) $F_{X,Y}(x,y) = \int_0^y\int_0^x (x+y)\,dx\,dy = \int_0^y \left[\frac12x^2+yx\right]_0^x dy$

$=\int_0^y\left[\frac12x^2+yx\right]dy = \left[\frac12x^2y+\frac12y^2x\right]_0^y = \boxed{\frac12x^2y+\frac12y^2x = F_{X,Y}(x,y) \quad 0\le x,y\le 1}$

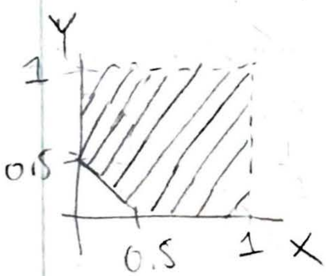
3) $f_X(x) = \int_0^1[x+y]\,dy = \left[xy+\frac12y^2\right]_{0=y}^1 = \boxed{x+\frac12 = f_X(x)}$

$f_Y(y) = \int_0^1[x+y]\,dx = \left[\frac12x^2+yx\right]_{0=x}^1 = \boxed{\frac12+y=f_Y(y)}$

4) $P(Y<X^2) = \iint_{x^2>y} f_{XY}(x,y)\,dx\,dy = \int_0^1\int_{y^{1/2}=x}^1 x+y\,dx\,dy$

$=\int_0^1\left[\frac12x^2+yx\right]_{x=y^{1/2}}^1 dy = \int_0^1 \frac12 + \cancel{y - \frac12 y} - y^{3/2}\,dy = \left[\frac12 y + \frac14 y^2 - \frac25 y^{5/2}\right]_0^1$

$= \frac12 + \frac14 - \frac23 = \boxed{\frac{7}{20} = 0.35 = P(Y<X^2)}$



$P[X+Y>0.5] = 1 - P[X+Y\le 0.5]$, $P[X+Y\le 0.5]$

$=\int_0^{1/2}\int_0^{\frac12-x}(x+y)\,dy\,dx = \int_0^{1/2}\left[xy+\frac12y^2\right]_0^{\frac12-x=y}dx =$

$(\frac12-x)(\frac12-x)$
$\frac14 - x + x^2$

$\int_0^{1/2} x(\frac12-x)+\frac12(\frac12-x)^2\,dx = \int_0^{1/2} \cancel{\frac12 x} - x^2 + \frac18 - \cancel{\frac12 x} + \frac12 x^2\,dx$

$\left[-\frac13x^3+\frac18x+\frac16x^3\right]_0^{1/2} = -\frac1{24}+\frac1{16}+\frac1{48} = \frac1{24} \therefore 1-\frac1{24}=\frac{23}{24}$

$\boxed{P[X+Y>0.5] = \frac{23}{24} = 0.958}$

3. $\int_0^{10} 0.3\,dx = 3$. Definetly not a pdf since we know that for $0 \le x \le 10$, $P(x) = 0.3$, this means that the total probability exceeds 1 since $3 > 1$. It is not possible to have $P(x) < 0$ so we cannot go back down to 1.

4. 1) $P(A \text{ and } B) = P(A)P(B) = 0.3 \times 0.6 = \boxed{\frac{9}{50} = 0.18 = P(A \times B)}$

$P(A \text{ or } B) = P(A) + P(B) = 0.3 + 0.6 = \boxed{0.9 = P(A \text{ or } B)}$

$P(A|B) = P(A) = \boxed{0.3 = P(A|B)}$

5. P is possibly a probability function since $\sum_{s \in \Omega} P(s) = 1$ and thus that means that $P(C) + P(D) = 0.1$ so it's possible so long as C & D meet that constraint.

6. Optimal $W = [-0.1549, 5.0094, -1.8536, 1.6013]$
Error: 0.8599

For part 2, regularization does not improve results since we are dealing with a linear regression problem (Plane in N-space) so no overfitting occurs. moreover, there's enough training data, that the problem is not ill-posed. From the results adding $\lambda \sim [0, 1]$ lead to a higher error since it serves no purpose as expected. So the optimal λ occurs at $\lambda = 0$ since this will give the minimum error.