homework sheet 03

Marco Seravalli
03626387
`marco.seravalli@tum.de`

Nikola Tchipev
03625168
`n.tchipev@tum.de`

1 Basic Probability

Problem 1

Probability to be a terrorist $P(x=T)=\frac{1}{100}$
Probability to be a citizen $P(x=C)=\frac{99}{100}$
Probability to recognize a terrorist as such $P(T|x=T)=\frac{95}{100}$
Probability to recognize a citizen as terrorist $P(T|x=C)=1-P(C|x=C)=\frac{5}{100}$
Probability that my neighbor is actually a terrorist

$$P(x=T|T)=\frac{P(x=T)P(T|x=T)}{P(x=T)P(T|x=T)+P(x=C)P(T|x=C)}$$

$$P(x=T|T)=\frac{\frac{1}{100}\frac{95}{100}}{\frac{1}{100}\frac{95}{100}+\frac{99}{100}\frac{5}{100}}=\frac{19}{118}\approx 0.16$$

Problem 2

Notation $P(2d=RR)$ probability to draw two red balls, $P(RR)$ probability that the box contains two red balls.
Probability of having 2 red balls in the box $P(RR)=\frac{1}{4}$
Probability of having one red and one white ball in the box $P(WR)=\frac{1}{2}$
Probability of having 2 white balls in the box $P(WW)=\frac{1}{4}$
Probability that all the balls in the box are red given that 3 red balls were extracted in 3 different draws:

$$P(RR|3d=RRR)=$$

$$=\frac{P(1d=R|RR)^3P(RR)}{P(1d=R|RR)^3P(RR)+P(1d=R|WR)^3P(WR)+P(1d=R|WW)^3P(WW)}$$

We know that

$$P(1d=R|WW)=0$$

$$P(1d=R|WR)=\frac{1}{2}$$

$$P(1d=R|RR)=1$$

Hence

$$P(RR|3d=RRR)=\frac{P(RR)}{P(RR)+P(1d=R|WR)^3P(WR)}=\frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{8}\frac{1}{2}}=\frac{4}{5}$$

Problem 3

Expected value:

$$E\left[X\right]=\int_{-\infty}^{+\infty}xf(x)dx=\int_{-\infty}^{0}xf(x)dx+\int_{0}^{1}xf(x)dx+\int_{1}^{+\infty}xf(x)dx$$

$$E\left[X\right]=0+\int_0^1 xdx+0=\left|\frac{x^2}{2}\right|_0^1=\frac{1}{2}$$

Variance:

$$V\left[X\right]=E\left[(X-E\left[X\right])^2\right]=\int_0^1\left(x-\frac{1}{2}\right)^2dx$$

$$V\left[X\right]=\int_0^1 x^2-x+\frac{1}{4}dx=\left|\frac{x^3}{3}-\frac{x^2}{2}+\frac{x}{4}\right|_0^1=\frac{1}{3}-\frac{1}{2}+\frac{1}{4}=\frac{1}{12}$$

Problem 4

Part A

Prove

$$E\left[X\right] = E_Y\left[E_{X|Y}\left[X\right]\right]$$

We know that

$$E\left[X\right] = \sum_x x p_X(x)$$

$$E\left[X\right] = \sum_x x \sum_y p_{XY}(x,y)$$

$$E\left[X\right] = \sum_x x \sum_y \frac{p_{XY}(x,y)}{p_Y(y)} p_Y(y)$$

$$E\left[X\right] = \sum_y \sum_x x \frac{p_{XY}(x,y)}{p_Y(y)} p_Y(y)$$

Use Bayes rule

$$E\left[X\right] = \sum_y \sum_x x p_{X|Y}(x|y) p_Y(y)$$

$$E\left[X\right] = \sum_y E_{X|Y}\left[X\right] p_Y(y)$$

$$E\left[X\right] = E_y\left[E_{X|Y}\left[X\right]\right]$$

Part B

Prove

$$Var_X\left[X\right] = E_Y\left[Var_{X|Y}\left[X\right]\right] + Var_Y\left[E_{X|Y}\left[X\right]\right]$$

The R.H.S. can be also rewritten as

$$E_Y\left[Var_{X|Y}\left[X\right]\right] + Var_Y\left[E_{X|Y}\left[X\right]\right] =$$

$$= E_Y\left[E_{X|Y}\left[X^2\right] - E_{X|Y}\left[X\right]^2\right] + E_Y\left[E_{X|Y}\left[X\right]^2\right] - E_Y\left[E_{X|Y}\left[X\right]\right]^2$$

By exploiting the linearity of E we get

$$E_Y\left[Var_{X|Y}\left[X\right]\right] + Var_Y\left[E_{X|Y}\left[X\right]\right] =$$

$$= E_Y\left[E_{X|Y}\left[X^2\right]\right] - E_Y\left[E_{X|Y}\left[X\right]^2\right] + E_Y\left[E_{X|Y}\left[X\right]^2\right] - E_Y\left[E_{X|Y}\left[X\right]\right]^2$$

$$= E_Y\left[E_{X|Y}\left[X^2\right]\right] - E_Y\left[E_{X|Y}\left[X\right]\right]^2$$

If the result of the Part A is reused we obtain

$$E_Y\left[Var_{X|Y}\left[X\right]\right] + Var_Y\left[E_{X|Y}\left[X\right]\right] = E_X\left[X^2\right] - E_X\left[X\right]^2 = Var_X\left[X\right]$$

2 Probability Inequalities

2.1 Markov Inequality

Let $I_{X>c}$ be the indicator random variable of the event $X > c$:

$$I_{X>c} = \begin{cases} 1 & \text{if } X > c, \\ 0 & \text{otherwise.} \end{cases}$$

We then have

$$cI_{X>c} = \begin{cases} c & \text{if } X > c, \\ 0 & \text{otherwise,} \end{cases}$$

and hence

$$cI_{X>c} \leq X,$$

since X is non-negative. Now, by the monotonicity of the expectation operator we have

$$E[cI_{X>c}] \leq E[X].$$

Taking linearity of E, into account, we get

$$cE[I_{X>c}] \leq E[X].$$

Note that we can compute $E[I_{X>c}]$:

$$\begin{aligned} E[I_{X>c}] &= \sum_x p(x) I_{X>c} \\ &= \sum_{x \leq c} p(x) I_{X>c} + \sum_{x>c} p(x) I_{X>c} \\ &= \sum_{x \leq c} p(x)0 + \sum_{x>c} p(x)1 \\ &= \sum_{x>c} p(x) \\ &= P(X > x). \end{aligned}$$

Thus, we arrive at

$$cP(X > x) = cE[I_{X>c}] \leq E[X],$$

and hence

$$P(X > x) \leq \frac{E[X]}{c}.$$

For the second part, let X be the random variable, describing the number of "heads" out of n coin tosses. Since we are throwing a fair coin, we clearly have

$$E[X] = \frac{n}{2}.$$

Now we have $c = \frac{3}{4}n$ and by the Markov inequality we obtain

$$P\left(X > \frac{3}{4}n\right) \leq \frac{\frac{n}{2}}{\frac{3}{4}n} = \frac{2}{3}$$

and

$$P\left(X > \frac{3}{4}n\right) \leq \frac{2}{3}.$$

2.2 Chebyshev Inequality

Similarly, let $I_{(X-E[X])^2>c^2}$ be the indicator random variable of the event $(X-E[X])^2 > c^2$:

$$I_{(X-E[X])^2>c^2} = \begin{cases} 1 & \text{if } (X-E[X])^2 > c^2, \\ 0 & \text{otherwise.} \end{cases}$$

We then have

$$c^2 I_{(X-E[X])^2>c^2} = \begin{cases} c^2 & \text{if } (X-E[X])^2 > c^2, \\ 0 & \text{otherwise,} \end{cases}$$

and hence

$$c^2 I_{(X-E[X])^2>c^2} \leq (X-E[X])^2.$$

Now, by the monotonicity of the expectation operator we have

$$E[c^2 I_{(X-E[X])^2>c^2}] \leq E[(X-E[X])^2] = Var[E].$$

Taking linearity of E, into account, we get

$$c^2 E[I_{(X-E[X])^2>c^2}] \leq Var[E].$$

Note that we can compute $E[I_{(X-E[X])^2>c^2}]$:

$$\begin{aligned} E[I_{(X-E[X])^2>c^2}] &= \sum_x p(x) I_{(X-E[X])^2>c^2} \\ &= P\left((X-E[X])^2 > c^2\right) \\ &= P\left(|X-E[X]| > c\right) \end{aligned}$$

Thus, we arrive at

$$P(|X-E[X]| > c) \leq \frac{Var[X]}{c^2}.$$

For the second part, let again X be the random variable, describing the number of "heads" out of n coin tosses. Now, due to the symmetry of the problem and the fact that $E[X] = \frac{n}{2}$, we make the following observation:

$$P\left(X > \frac{3}{4}n\right) = P\left(X < \frac{1}{4}n\right).$$

Thus,

$$\begin{aligned} P\left(X > \frac{3}{4}n\right) &= \frac{2P\left(X > \frac{3}{4}n\right)}{2} & (1) \\ &= \frac{P\left(X > \frac{3}{4}n\right) + P\left(X < \frac{1}{4}n\right)}{2} & (2) \\ &= \frac{P\left(\{X > \frac{3}{4}n\} \cup \{X < \frac{1}{4}n\}\right)}{2} & (3) \\ &= \frac{P\left(|X-E[X]| > \frac{1}{4}n\right)}{2}. & (4) \end{aligned}$$

Now, after applying Chebyshev's inequality, we obtain:

$$P\left(X > \frac{3}{4}n\right) \leq \frac{Var[X]}{2\left(\frac{1}{4}n\right)^2} = \frac{8Var[X]}{n^2}.$$

Now we just need the variance:

$$\begin{aligned} Var[X] &= E[(X-E[X])^2] \\ &= \sum_x p(x)\left(x - \frac{n}{2}\right)^2 \\ &= \sum_{k=0}^{n} \frac{\binom{n}{k}}{2^n}\left(k - \frac{n}{2}\right)^2. \end{aligned}$$

2.3 Jensen's Inequality

The statement is obviously true for $n = 1$:

$$f(1 \cdot x_1) \leq 1 \cdot f(x_1)$$

and for $n = 2$ since f is convex:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2),$$

which gives us the base step of the induction.

Now suppose it is true for $n = k$ and consider $n = k+1$:

$$\begin{aligned}
f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) &= f\left(\sum_{i=1}^{k} \lambda_i x_i + \lambda_{k+1} x_{k+1}\right) && (5)\\
&= f\left((1-\lambda_{k+1})\left(\frac{\sum_{i=1}^{k} \lambda_i x_i}{1-\lambda_{k+1}}\right) + \lambda_{k+1} x_{k+1}\right) && (6)\\
&\leq (1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i x_i}{1-\lambda_{k+1}}\right) + \lambda_{k+1} f(x_{k+1}) && (7)\\
&\leq (1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i\right) + \lambda_{k+1} f(x_{k+1}) && (8)\\
&\leq (1-\lambda_{k+1}) \left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} f(x_i)\right) + \lambda_{k+1} f(x_{k+1}) && (9)\\
&= \sum_{i=1}^{k} \lambda_i f(x_i) + \lambda_{k+1} f(x_{k+1}) && (10)\\
&= \sum_{i=1}^{k+1} \lambda_i f(x_i). && (11)
\end{aligned}$$

In going from (6) to (7) we used the definition of convexity of f. In going from (8) to (9) we use the fact that

$$\sum_{i=1}^{k} \lambda_i = 1 - \lambda_{k+1}, \quad \text{hence} \quad \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = 1$$

and apply the inductive hypothesis for $n = k$.