Exercise 1

Machine Learning I

1A-1.

Example of mixed distribution: You are throwing n balls into the air. Let $P(X = i)$ be the probability, that the i-th ball hits the ground first and $f_Y(i)$ be the density of its exact displacement during the fall. The joint distribution $P(X = i, a < Y < b)$ expects X to have finite and Y uncountably infinite many realizations.

1A-2.

Counterexample:

Given X, let A, B be the results of two successive *independent* coin tosses.

Let $C = B$.

We now have:

$$P(A, B|X) = \underbrace{P(A|X) \cdot P(B|X)}_{conditional\ independence\ A,B}$$

$$P(A, C|X) = P(A, B|X) = \underbrace{P(A|X) \cdot P(C|X)}_{conditional\ independence\ A,B} .$$

But:

$$P(B, C|X) \underbrace{\neq}_{C=B} P(B|X) \cdot P(C|X). \square$$

This can be further analyzed.

We have the following constraints:

Given the above restrictions R, we analyze the validity of the relationship:

$$R \Longrightarrow P(B, C|X) = P(B|X) \cdot P(C|X)$$

Solution: Transform the joint probability $\mathrm{P}(B, C|X)$ into $P(B, C|X) = P(C|B, X) \cdot P(B|X)$.

Excluding the degenerate case $P(B|X) = 0$, this leads to:

$$
\begin{aligned}
& P(B,C|X) = P(B|X)\cdot P(C|X)\\
\Leftrightarrow\quad & P(C|B,X)\cdot P(B|X) = P(B|X)\cdot P(C|X)\\
\Leftrightarrow\quad & P(C|B,X) = P(C|X)\\
\Leftrightarrow\quad & \frac{P(C,B,X)}{P(B,X)} = P(C|X).
\end{aligned}
$$

The above only holds iff $P(C|B,X) = P(C|X)$, which is generally not true, so conditional independence of A,B and A,C is not sufficient for transitivity.

Just imagine: If $P(C|B,X) = P(C|X)$ were always valid, we could not learn about C by gathering data B,X. In other words: More data would not make our estimates any better. This runs counter to most situations.

1A-3.

Definition of the events:
$E = Person\ is\ guilty,$
$T = Person\ passes\ the\ test.$

(i) The negations $\bar{T},\bar{E}$ can be read as *not*.

$$P(E|\bar{T}) = \frac{P(\bar{T}|E)\cdot P(E)}{P(\bar{T})} = \frac{\frac{5}{6}\cdot\frac{1}{3}}{\frac{7}{18}} = \frac{5}{7},$$

with

$$P(\bar{T}|E) = \frac{5}{6},$$

$$P(E) = \frac{1}{3},$$

$$P(\bar{T}) = P(E)\cdot P(\bar{T}|E) + P(\bar{E})\cdot P(\bar{T}|\bar{E}) = \frac{1}{3}\cdot\frac{5}{6} + \frac{2}{3}\cdot\frac{1}{6} = \frac{7}{18}$$

(ii)

$$P(E|\bar{T},\bar{T}) = \frac{P(\bar{T},\bar{T}|E)}{P(\bar{T},\bar{T})} = \underbrace{\frac{P(\bar{T}|E)\cdot P(\bar{T}|E)\cdot P(E)}{P(\bar{T},\bar{T})}}_{conditional\ independence} = \frac{\left(\frac{5}{6}\right)^2\cdot\frac{1}{3}}{\frac{1}{3}\cdot\left(\frac{5}{6}\right)^2 + \frac{2}{3}\cdot\left(\frac{1}{6}\right)^2} = 0.\overline{925}.$$

Using the conditional independence of $P(\bar{T},\bar{T}|E)$ and independence of testing $P(\bar{T},\bar{T}|E)$.

1A-4.

$$E[X] = \sum_{i=1}^{6} i\cdot P(X=i) = (1+2+3)\cdot\frac{1}{12} + (4+5)\cdot\frac{1}{6} + 6\cdot\frac{5}{12} = 4.5.$$

$$Var[X] = E[X^2] - E[X]^2 = \left[(1+4+9)\cdot\frac{1}{12} + (16+25)\cdot\frac{1}{6} + 36\cdot\frac{5}{12}\right] - 4.5^2 = \frac{276}{12} - 4.5^2$$
$$= 2.75\,.$$

$$E[X_1 + E_2] = 2 \cdot E[X_1] = 9.$$

1A-5.

$$\begin{aligned} E[(X - E[X])(Y - E[Y])] &= E\big[XY - XE[Y] - YE[X] + E[X]E[Y]\big] \\ &= E[XY] - E[Y]E[X] - E[X]E[Y] + E[X]E[Y] \\ &= \underbrace{E[XY] - E[X]E[Y]}_{*} \\ &= 0. \end{aligned}$$

$$E[XY] = \int_{y_0}^{y_1}\int_{x_0}^{x_1} xy \cdot f_{X,Y}(x,y)dx\,dy = \int_{y_0}^{y_1} y \int_{x_0}^{x_1} x \cdot \underbrace{f_X(x)f_y(y)}_{independence}\,dx\,dy$$
$$= \int_{y_0}^{y_1} yf_y(y)\int_{x_0}^{x_1} x\,f_X(x)dx\,dy = \int_{y_0}^{y_1} yf_y(y)\cdot E[X]\,dy = E[X]\int_{y_0}^{y_1} yf_y(y)\,dy$$
$$= E[X]E[Y].$$

Utilizing that $*$ $E[XY] = E[X]E[Y]$ if X, Y are independent.

If X, Y are discrete equivalent steps can be taken to prove the conjecture.