Probability Theory exam
December 20th, 2022

The. 1. a) E, F events: $\mathbb{P}(E|F) = \dfrac{\mathbb{P}(E\cap F)}{\mathbb{P}(F)}$

b) $E_1, E_2, \dots$ form a complete event system in Ω if $i \neq j$: $E_i \cap E_j = \emptyset$ and $\bigcup_i E_i = \Omega$

c) $\boxed{\text{Thm}}$: Let F be an event and $E_1, E_2, \dots$ be a complete event system in Ω. If $\forall i$: $\mathbb{P}(E_i) > 0$, then

$$\mathbb{P}(F) = \sum_i \mathbb{P}(F|E_i)\,\mathbb{P}(E_i)$$

proof: $\mathbb{P}(F) = \mathbb{P}(F\cap\Omega) = \mathbb{P}(F\cap(\bigcup_i E_i)) =$

$$= \mathbb{P}\Big(\bigcup_i (F\cap E_i)\Big) \underset{\text{the union is disjoint}}{=} \sum_i \mathbb{P}(F\cap E_i) = \sum_i \underbrace{\frac{\mathbb{P}(F\cap E_i)}{\mathbb{P}(E_i)}}_{=\mathbb{P}(F|E_i)}\mathbb{P}(E_i) \qquad \square$$

d) $\boxed{\text{Thm}}$: Let F be an event st. $\mathbb{P}(F) > 0$. Let $E_1, E_2, \dots$ be a complete event system st. $\forall i$: $\mathbb{P}(E_i) > 0$. Then

$$\forall i: \mathbb{P}(E_i|F) = \frac{\mathbb{P}(F|E_i)\,\mathbb{P}(E_i)}{\sum_i \mathbb{P}(F|E_i)\,\mathbb{P}(E_i)}$$

proof: $\mathbb{P}(E_i|F) = \dfrac{\mathbb{P}(E_i\cap F)}{\mathbb{P}(F)} = \dfrac{\mathbb{P}(F|E_i)\,\mathbb{P}(E_i)}{\sum_i \mathbb{P}(F|E_i)\,\mathbb{P}(E_i)}$ ← def. in part a) ← thm in part c) $\square$

The. 2. a) $X \sim Bin(n_1, p)$, $Y \sim Bin(n_2, p)$ independent

claim: $X+Y \sim Bin(n_1+n_2, p)$

proof: $\mathbb{P}(X+Y=k) = \sum_{i=0}^{k} \mathbb{P}(X=k-i)\,\mathbb{P}(Y=i) = \sum_{i=0}^{k} \binom{n_1}{k-i} p^{k-i}(1-p)^{n_1-k+i} \binom{n_2}{i} p^i (1-p)^{n_2-i} =$

$$= p^k (1-p)^{n_1+n_2-k} \underbrace{\sum_{i=0}^{k} \binom{n_1}{k-i}\binom{n_2}{i}}_{=\binom{n_1+n_2}{k}} = \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k} \qquad \square$$

← number of different k member groups in a population of n_1+n_2 people

b) Let $Z \sim Bin(n, p)$ and $Z_1, \dots, Z_n \sim Ber(p)$ independent.

$Z = \sum_{i=1}^{n} Z_i \Rightarrow \mathbb{E}[Z] = \mathbb{E}[\sum_{i=1}^{n} Z_i] = \sum_{i=1}^{n} \mathbb{E}[Z_i] = np$

$Var(Z) = Var(\sum_{i=1}^{n} Z_i) = \sum_{i=1}^{n} \underbrace{Var(Z_i)}_{=\mathbb{E}[Z_i^2] - \mathbb{E}^2[Z_i] = p - p^2 = p(1-p)} = np(1-p) \Rightarrow D[Z] = \sqrt{np(1-p)}$

c) $U \sim Bin(n, 1/2)$, $V \sim Bin(n, 2/3)$ independent

n is really big $\Rightarrow$ de Moivre-Laplace

$$\frac{U - n/2}{\sqrt{n/4}} =: \hat{U} \sim N(0,1)$$

$$\frac{V - \frac{2n}{3}}{\sqrt{2n/9}} =: \hat{V} \sim N(0,1)$$

$$\mathbb{P}(4U - 3V \geq \sqrt{n}) = \mathbb{P}\Bigg(4 \cdot \underbrace{\frac{U - n/2}{\sqrt{n/4}}}_{\hat{U}} \cdot \frac{\sqrt{n/4}}{\underbrace{\sqrt{16 \cdot n/4 + 9 \cdot 2n/9}}_{=\sqrt{6n}}} - 3 \cdot \underbrace{\frac{V - 2n/3}{\sqrt{2n/9}}}_{\hat{V}} \cdot \frac{\sqrt{2n/9}}{\sqrt{6n}} \geq \frac{\sqrt{n} \cancel{- 4\frac{n}{2} + 3\frac{2n}{3}}}{\sqrt{6n}}\Bigg)$$

$$= \mathbb{P}\Bigg(\underbrace{4 \cdot \hat{U} \frac{\sqrt{n/4}}{\sqrt{6n}} - 3 \hat{V} \cdot \frac{\sqrt{2n/9}}{\sqrt{6n}}}_{\sim N(0,1), \text{ this is why we divided by } \sqrt{6n}} \geq \frac{1}{\sqrt{6}}\Bigg) = 1 - \underbrace{\Phi\left(\frac{1}{\sqrt{6}}\right)}_{\approx 0{,}6591} = \underline{0{,}3409}$$

Thm. 3. a) Markov's inequality: Let X be a non-negative random variable. Then for all $a > 0$: $\mathbb{P}(X > a) \leq \mathbb{E}[X]/a$

proof: $\mathbb{E}[X] \geq \mathbb{E}[X \cdot \mathbb{1}\{X > a\}] \geq \mathbb{E}[a \mathbb{1}\{X > a\}] = a\,\mathbb{P}(X > a)$ $\square$

b) Chebisev's inequality: Let X be a random variable with $\mathbb{E}[X] = \mu$, $Var(X) = \sigma^2$. Then for all $b > 0$: $\mathbb{P}(|X - \mu| > b) \leq \sigma^2/b^2$

proof: $$\mathbb{P}(|X - \mu| > b) = \mathbb{P}((X - \mu)^2 > b^2) \underset{\text{(part a)}}{\leq} \frac{\mathbb{E}[(X - \mu)^2]}{b^2} = \frac{\sigma^2}{b^2} \quad \square$$

c) Weak law of large numbers: Let $X_1, X_2, \ldots$ be a sequence of iid random variables with expected value $\mathbb{E}[X_1] = \mu < \infty$ and variance $Var(X_1) = \sigma^2 < \infty$. Then

$$\forall \varepsilon > 0: \lim_{n \to \infty} \mathbb{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right) = 0$$

proof: Let $\overline{X}_n := \frac{X_1 + \ldots + X_n}{n}$, then $\mathbb{E}[\overline{X}_n] = \mu$, $Var(\overline{X}_n) = \sigma^2/n$

$$\mathbb{P}(|\overline{X}_n - \mu| > \varepsilon) \underset{\text{(part b)}}{\leq} \frac{Var(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n \varepsilon^2} \xrightarrow[n \to \infty]{} 0 \quad \square$$

Prac. 1. There are a lot of lights on the tree and they burn out with a small probability.
$\Rightarrow$ We can model the number of burned out lights with Poisson distribution
$\Rightarrow$ They burn out with the same probability independently of each other, and this probability doesn't change in time, so we have a Poisson process.

$N(t)$:= number of lights that burned out in t days, $\mathbb{E}[N(7)] = 3 \Rightarrow \underline{\lambda = \frac{3}{7}}$ is the parameter of the process

a) $$\mathbb{P}(N(3) \geq 2) = 1 - \mathbb{P}(N(3) = 0) - \mathbb{P}(N(3) = 1) = \underline{1 - e^{-9/7} - \tfrac{9}{7} e^{-9/7}} \approx 0{,}3681$$

Sunday 8 am — Monday 8 am — Tuesday 8 am — Wednesday 8 am

b) $N(t)$ is a Poisson process with parameter $\lambda = 3/7$

$\Rightarrow$ the waiting time between 2 points has distribution $Exp(3/7)$

T = time until the next light burns out

$\xrightarrow{\text{memorylessness}}$ $T \sim Exp(3/7)$, $E[T] = 7/3$ days $=$ 56 hours

c) First, let's count the expected number of days until he replaces a light again.

X = number of days until the next light burns out

$X \sim Geo(p)$, where $1-p = P(N(1)=0) = e^{-3/7} \approx 0{,}65$

$E[X] = \frac{1}{p} \approx 2{,}87$ days $= 68{,}85$ hours

(as he only checks the lights once every day)

Prac. 2. a) X = the number rolled with a fair die

$E[X] = \frac{1}{6}\sum_{i=1}^{6} i = \frac{21}{6}$, $Var(X) = E[X^2] - E^2[X] = \frac{546}{36} - \frac{441}{36} = \frac{105}{36} \Rightarrow D[X] = \frac{\sqrt{105}}{6} \approx 1{,}708$

$E[X^2] = \frac{1}{6}\sum_{i=1}^{6} i^2 = \frac{91}{6}$

b) A = amount of money that Andrew "wins"

$A = (X_1+X_2+X_3+X_4) - (X_1+X_2-X_3-X_4)^2$

$E[A] = 4\cdot E[X_1] - E[(X_1+X_2-X_3-X_4)^2] = 4\cdot\frac{21}{6} - 4\cdot\frac{91}{6} + 4\left(\frac{21}{6}\right)^2 = 2{,}333 > 0$

$E[(X_1+X_2-X_3-X_4)^2] = 4\cdot E[X_1^2] + 4E[X_1X_2] - 8E[X_1X_2]$

they are iid

$\Rightarrow$ Andrew is favoured.

Prac. 3. (X,Y) jointly normal with $E[X]=E[Y]=0$ and $\underline{C} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

a) $Z = \alpha X + \beta Y \Rightarrow Z$ always has normal distribution (except $\alpha=0, \beta=0$)

$E[Z] = \alpha E[X] + \beta E[Y] = 0$

$Var(Z) = Cov(\alpha X+\beta Y, \alpha X+\beta Y) = \alpha^2 Var(X) + \beta^2 Var(Y) + 2\alpha\beta\, Cov(X,Y) =$
$= 2\alpha^2 + 2\beta^2 + 2\alpha\beta$

$Cov(X,Z) = Cov(X, \alpha X+\beta Y) = \alpha Var(X) + \beta\, Cov(X,Y) = 2\alpha + \beta$

X, Z are independent $\Longleftrightarrow Cov(X,Z) = 0$
they are normal jointly

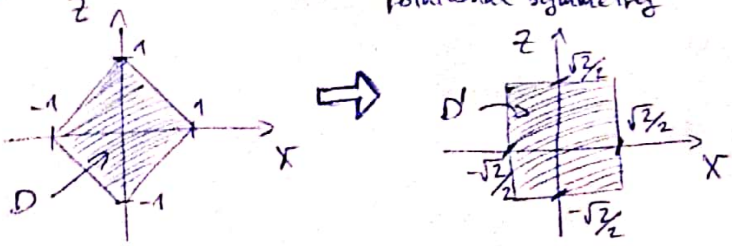
We want $\left.\begin{matrix} 2\alpha+\beta = 0 \\ 2\alpha^2+2\beta^2+2\alpha\beta = 2 \end{matrix}\right\} \xRightarrow[\alpha>0]{} \alpha = \frac{1}{\sqrt{3}}, \beta = -\frac{2}{\sqrt{3}}$

b) $f_{X,Z}(x,z) = \frac{1}{2\pi} e^{-\frac{x^2+z^2}{4}} = f_X(x) f_Z(z)$

$P(|X|+|Z|<1) = \iint_D f_{X,Z}(x,z)\,dx\,dz \underset{\text{rotational symmetry}}{=} \iint_{D'} f_{X,Z}(x,z)\,dx\,dz = P(-\sqrt{2}<X<\sqrt{2})P(-\sqrt{2}<Z<\sqrt{2}) =$

$= P(-\frac{1}{2} < \underbrace{\frac{X}{\sqrt{2}}}_{\sim N(0,1)} < \frac{1}{2})P(-\frac{1}{2} < \underbrace{\frac{Z}{\sqrt{2}}}_{\sim N(0,1)} < \frac{1}{2}) =$

$= (\underbrace{2\Phi(\tfrac{1}{2}) - 1}_{=0{,}383})^2 \approx 0{,}1467$

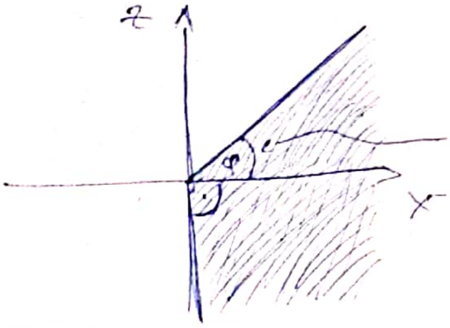


BONUS/ $Z = \frac{1}{\sqrt{3}}X - \frac{2}{\sqrt{3}}Y \Rightarrow Y = \frac{1}{2}X - \frac{\sqrt{3}}{2}Z$

$\mathbb{P}(X>0, Y>0) = \mathbb{P}(X>0, \frac{1}{2}X - \frac{\sqrt{3}}{2}Z > 0) = \mathbb{P}(X>0, X>\sqrt{3}Z) =$

$= \mathbb{P}(X>0, \frac{1}{\sqrt{3}}X > Z) = \frac{120}{360} \left(= \frac{2\pi/3}{2\pi}\right) = \frac{1}{3}$

their angle is uniformly distributed



$\varphi = \text{arctg}\left(\frac{1}{\sqrt{3}}\right) = 30°$

Alternative solutions

Prac. 2. b) $\mathbb{E}[A] = 4 \cdot \mathbb{E}[X_1] - \underbrace{\mathbb{E}[(X_1+X_2-X_3-X_4)^2]}_{= Var(X_1+X_2-X_3-X_4)} = 4 \cdot \mathbb{E}[X_1] - 4 \cdot Var(X_1) = 14 - \frac{35}{3} > 0$

$\mathbb{E}[X_1+X_2-X_3-X_4] = 0$

The. 2. a) $X = \sum_{i=1}^{n_1} X_i$, $X_i \sim Ber(p)$ indep.

$Y = \sum_{j=1}^{n_2} Y_j$, $Y_j \sim Ber(p)$ indep.

$\Rightarrow X + Y = \sum_{i=1}^{n_1} X_i + \sum_{j=1}^{n_2} Y_j$ It is a sum of n_1+n_2 many independent $Ber(p)$ random variables

$\Rightarrow X+Y \sim Bin(n_1+n_2, p)$