1

a.) $X \sim Bin(n=6, p=0.05) \Rightarrow \boxed{g_X(z) = (q+pz)^n = (0.95+0.05z)^6}$

$= \left(\frac{19+z}{20}\right)^6$

b.) $Y \sim Bin(n=60, p=0.05) \Rightarrow \boxed{g_Y(z) = (0.95+0.05z)^{60} = \left(\frac{19+z}{20}\right)^{60}}$

c.) Legyen N a megvett alma-csomagok száma, vagyis az első szétnézés előtt eltelt napok száma.

Így $N \sim PesszGeom(p)$ ahol $p_N = \frac{1}{3}$ a szétnézés valószínűsége.

Legyen $i = 1, 2, 3, \ldots$-ra X_i az i-edik csomagban lévő kukacos almák száma: $X_i \sim Bin(n=6, p_X = 0.05)$.

Ezzel az összes megvett kukacos alma száma

$Z = \sum_{i=1}^{N} X_i$ véletlen tagszámú összeg

$\Rightarrow \boxed{\mathbb{E}Z = \mathbb{E}N\,\mathbb{E}X_i = \left(\frac{1}{p_N} - 1\right)\cdot(n\,p_X) = \left(\frac{1}{1/3} - 1\right)\cdot(6 \cdot 0.05) = 2 \cdot \frac{6}{20} = 0.6}$

② Azok a pillanatok, amikor Mérickánál eszébe jut, hogy fel kéne csavarni a fűtést, Poisson pontfolyamatot alkotnak, $\lambda = 2$ intenzitással (ha az időt órában mérjük), mert
- az események közötti idők független, azonos, exponenciális eloszlásúak
- a követési idő átlag $\frac{1}{2}$ óra.

Ezen belül azok a pillanatok, amikor tényleg feljebb veszi, ill. amikor mégsem, az előző Poisson folyamat ~~ritkításai~~ szinezései $p = \frac{1}{6}$ ill. $q = \frac{5}{6}$ súlyokkal. Ezért

a.) Legyen X: ahányszor $t = 2$ óra alatt feljebb vette a fűtést. $X \sim Poi(\lambda \cdot t \cdot p) = Poi(2 \cdot 2 \cdot \frac{1}{6}) = Poi(\frac{2}{3})$

$$\Rightarrow \boxed{\mathbb{P}(X=0) = e^{-2/3} \approx 0{,}51 = 51\%}$$

b) Legyen Y: ahányszor a $t \in [2,3]$ ($\Delta t = 1$ hosszú) ~~inter~~ időintervallumban feljebb vette a fűtést.

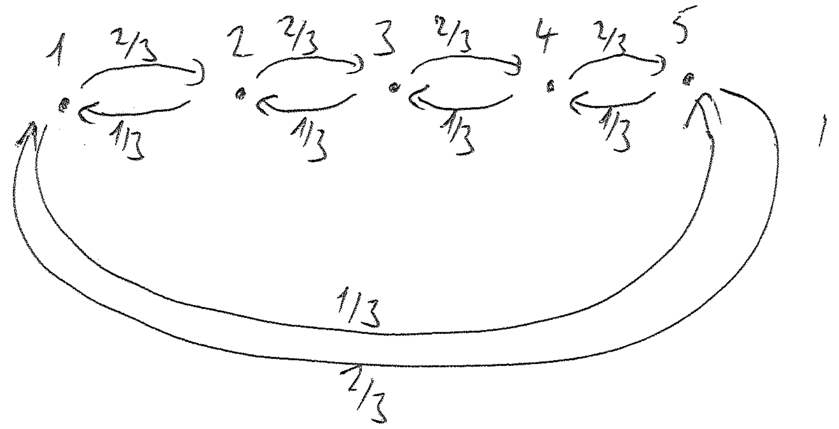
$Y \sim Poi(\lambda \cdot \Delta t \cdot p) = Poi(2 \cdot 1 \cdot \frac{1}{6}) = Poi(\frac{1}{3})$ és független X-től

$$\Rightarrow \boxed{\mathbb{P}(Y \geq 2 | X=0) = \mathbb{P}(Y \geq 2) = 1 - [\mathbb{P}(Y=0) + \mathbb{P}(Y=1)] = 1 - e^{-\frac{1}{3}}(1 + \frac{1}{3}) = 1 - \frac{4}{3} e^{-1/3} \approx 0{,}045 = 4{,}5\%}$$

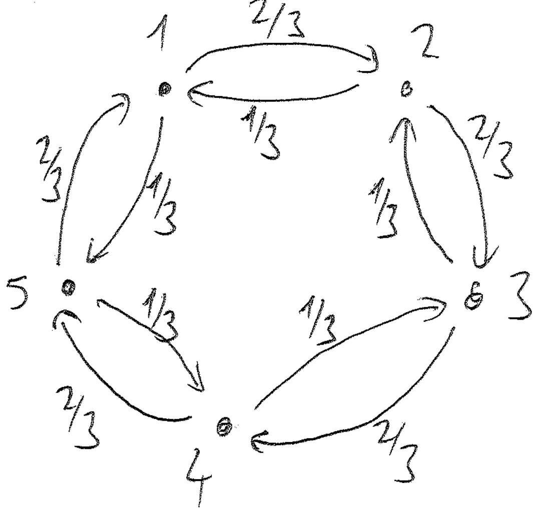
c.) Legyen Z: ahányszor 2 óra alatt eszébe jutott, de aztán mégse vette feljebb. $Z \sim Poi(\lambda \cdot t \cdot q) = Poi(2 \cdot 2 \cdot \frac{5}{6}) = Poi(\frac{10}{3})$ és független X-től

$$\Rightarrow \boxed{\mathbb{P}(Z=0 | X=0) = \mathbb{P}(Z=0) = e^{-10/3} \approx 0{,}036 = 3{,}6\%}$$

③ legyen $X_n \in S = \{1,2,3,4,5\}$, hogy hanyadik feladatnál jár n ugrás után. X_n diszkrét idejű, időben homogén Markov lánc, gráf-reprezentációja



de körbe lerajzolva szebb:



Az átmenetmátrix $P = \begin{pmatrix} 0 & 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 2/3 & 0 & 0 \\ 0 & 1/3 & 0 & 2/3 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 2/3 & 0 & 0 & 1/3 & 0 \end{pmatrix}$

és $X_0 = 1$.

③ – folytatás –

a.) $\mathbb{P}(X_4 = 1) = ?$ (miközben tudjuk, hogy $X_0 = 1$).

$X_4 = 1$ úgy eshet meg, ha a 4 lépésből pontosan 2-szer ugrik lefelé (pontosabban: lefelé VAGY 1-ből 5-be) és 2-szer felfelé (pontosabban: felfelé VAGY 5-ből 1-be)

Ez $\binom{4}{2} = 6$ féleképpen lehetséges.

[Gyengébbek kedvéért:

$1 \to 5 \to 4 \to 5 \to 1$, $1 \to 5 \to 1 \to 5 \to 1$,

$1 \to 5 \to 1 \to 2 \to 1$, $1 \to 2 \to 1 \to 5 \to 1$,

$1 \to 2 \to 1 \to 2 \to 1$, $1 \to 2 \to 3 \to 2 \to 1$]

Ezek mindegyikének valószínűsége $\left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^2$, tehát

$$\boxed{\mathbb{P}(X_4 = 1) = \binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2 = 6 \cdot \frac{2^2}{3^4} = \frac{24}{81} \approx 0.296 = 29.6\%}$$

b.) A Markov lánc véges állapotterű, irreducibilis (mert mindenhonnan mindenhova el lehet jutni) és aperiodikus (mert pl. 1-ből 1-be el lehet jutni 2 és 5 lépésben is).

A kérdés $\mathbb{P}(X_{30} = 1 \mid X_0 = 1)$. Mivel $n = 30$ hosszú idő, a Markov láncok alaptétele szerint $\mathbb{P}(X_{30} = 1 \mid X_0 = 1) \approx \pi_1$, ahol π az egyetlen stac. eloszlás.

A 2. rajz szimmetriájából látszik, hogy $\pi = \left(\frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\right)$ stacionárius $\Rightarrow$ $\boxed{\mathbb{P}(X_{30} = 1 \mid X_0 = 1) \approx \pi_1 = \frac{1}{5} = 20\%}$.

③ – 2. folytatás –

[Persze az, hogy a $\pi = (\frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\ \frac{1}{5}\ \frac{1}{5})$ egyenletes eloszlás stacionárius, abból is látszik, hogy a P mátrix bisztochasztikus (minden oszlopösszeg 1), vagy kiszámolható a $(P^T - \mathbb{1})\pi^T = 0$ lineáris egyenletrendszer megoldásával.]

c.) Ha 6 állapot lenne, akkor az X_n Markov lánc
BÓNUSZ
periodikus lenne $d=2$ periódussal, mert pl 1-ből 1-be csak páros sok lépésben lehetne visszajutni. Ekkor a $\pi = (\frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6})$ egyenletes eloszlás továbbra is stacionárius, de

$\mathbb{P}(X_{30} = 1 | X_0 = 1) \approx 2\pi_1 = \frac{2}{6} = \frac{1}{3}$, mert így, hogy a Markov lánc csak minden páros lépésben lehet 1-ben, ilyenkor kénytelen dupla val. széggel ott lenni.

(4) Az $n=10$ elemű mintából a likelihood-függvény

$$L(\theta)=\prod_{i=1}^{n} p_\theta(x_i) \quad \text{ahol} \quad p_\theta(x):=\mathbb{P}(X=x)=(x+1)(1-\theta)^x\theta^2$$

$x=0,1,2,\dots$

$$=\prod_{i=1}^{n}(x_i+1)(1-\theta)^{x_i}\theta^2.$$

Ebből a log-likelihood függvény

$$\ell(\theta)=\ln L(\theta)=\sum_{i=1}^{n}\left[\ln(x_i+1)+x_i\ln(1-\theta)+2\ln\theta\right]$$

Ennek maximuma csak ott lehet, ahol

$$\frac{d\ell}{d\theta}=\sum_{i=1}^{n}\left[0+x_i\frac{1}{1-\theta}\cdot(-1)+2\cdot\frac{1}{\theta}\right]=\frac{-1}{1-\theta}\sum_{i=1}^{n}x_i+\frac{2n}{\theta}=0,$$

Vagyis $\frac{2n}{\theta}=\frac{\sum_{i=1}^{n}x_i}{1-\theta}$.

Ezt megoldva
$$2(1-\theta)=\theta\frac{\sum_{i=1}^{n}x_i}{n}=\theta\bar{x}$$
$$2-2\theta=\theta\bar{x}$$
$$2=\theta(\bar{x}+2)$$
$$\theta=\frac{2}{\bar{x}+2} \quad \text{az egyetlen lehetséges maximumhely}$$

$\Rightarrow$ a maximum likelihood becslés

$$\boxed{\hat{\theta}_{ML}=\frac{2}{\bar{x}+2}}$$

esetünkben $\bar{x}=\frac{6+2+3+\dots+13}{10}=3.8$

$$\frac{2}{3.8+2}=\frac{2}{5.2}\approx 0.385$$

[Megjegyzés: X két független θ paraméterű pesszimista geometriai eloszlású val. változó összege — a kapott $\hat{\theta}_{ML}$ képlet ezzel konzisztens.]

(5) A szórás $\sigma = 0.25$ ismert, a null-hipotézis $H_0: m = \mu := 10$ egyenlőség az egyetlen minta várható értéke és egy hipotetikus érték között $\Rightarrow$ egymintás kétoldali u-próbát végzünk.

A teszt-statisztika a képletgyűjtemény szerint

$$u = \frac{\bar{x} - \mu}{\sigma}\sqrt{n} = \frac{\frac{101.1}{10} - 10}{0.25}\sqrt{10} = \frac{0.11}{0.25}\sqrt{10} \approx 1.39$$

A küszöbérték $K = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right)$ ahol $\varepsilon = 1 - 0.95 = 0.05$

mert 95% a konfidenciaszint

$$= \Phi^{-1}(0.975) \approx 1.96$$

Döntés: $|u| \leq K \Rightarrow$ a null-hipotézist elfogadjuk.