① Legyenek egy e-mail utódai mindazon e-mailek, amik a megválaszolását követő napon érkeznek. Legyen a legelső levél egymaga a „nulladik generáció", és $\forall n \geq 0$-ra az $(n+1)$-edik generáció álljon az n-edik generáció utódaiból. Végül legyen Z_n az n-edik generáció elemszáma.

A feladat szövege szerint Z_n Galton-Watson elágazó folyamat, $Z_0 = 1$ és az 1-lépéses utódszám-eloszlás

$X \sim \mathrm{Poi}(1)$

a.) A kérdés a folyamat kihalásának valószínűsége.

Mivel $m := \mathbb{E}X = 1$, a folyamat kritikus, így

$\boxed{\mathbb{P}(\text{kihalás}) = 1.}$ [A folyamat nem elfajult: $X \not\equiv 1$.]

b.) A kérdés a kihalásig érkező összes levelek számának várható értéke: $\mathbb{E}N = \mathbb{E}\sum_{n=0}^{\infty} X_n = ?$

Mivel a folyamat kritikus, $\boxed{\mathbb{E}N = \infty.}$

(2) A kiégési időpontok Poisson-pontfolyamatot alkotnak, $\lambda = 2$ $\left(\frac{\text{darab}}{\text{év}}\right)$ intenzitással, mivel $\frac{1}{\lambda} = \mathbb{E}(\text{köztes idő}) = \frac{1}{2}$ év.

a.) Legyen X a jövőre kiégő ~~égők~~ körték száma. $X \sim Poi(2)$,

így $\boxed{\mathbb{P}(X > 2)} = 1 - [\mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2)] = 1 - e^{-2}\left[\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!}\right] =$

$= 1 - e^{-2}(1+2+2) = \boxed{1 - 5e^{-2} \approx 0.32 = 32\%}$

b.) A Móricka és Pistike által cserélt körték időpontjai az eredeti Poisson-folyamat színezései, így maguk is ~~pontfolyamatok~~ Poisson-folyamatok és függetlenek.

Így ha ~~X~~ Y a Pistike által cserélt körték száma, Z pedig a Móricka által cseréltek száma, akkor

$Z \sim Poi(\frac{1}{2} \cdot 2) = Poi(1)$, és

$\boxed{\mathbb{P}(Z = 0 | Y = 3)} \overset{\text{függet-}}{\underset{\text{lenség}}{=}} \mathbb{P}(Z = 0) = e^{-1}\frac{1^0}{0!} = \boxed{e^{-1} \approx 0.37 = 37\%}$

(3) Legyen X_i az i-edik kérés kiszolgálásához szükséges idő, $\boxed{\text{ms-ban}}$.
$i=1,2,\dots,n$, ahol $n=10000$, és legyen
$S_n = X_1+\dots+X_n$. A kérdés a $\mathbb{P}(S_n \geq 20000)$ valószínűség CHT-becslésének hibája. Ez a Berry-Esseen tétel szerint

legfeljebb $\dfrac{C\delta}{\sqrt{n}\,\sigma^3}$, ahol

- $C = 0.4748$
- $n = 10000$
- σ a szórás; ehhez $m := EX = \frac{1}{10}\cdot 0 + \frac{8}{10}\cdot 1 + \frac{1}{10}\cdot 10 = 1.8$

$$\mathbb{E}(X^2) = \frac{1}{10}\cdot 0^2 + \frac{8}{10}\cdot 1^2 + \frac{1}{10}\cdot 10^2 = 10.8$$

$$\Rightarrow \sigma = \sqrt{E(X^2)-(EX)^2} = \sqrt{10.8-1.8^2} = \sqrt{7.56} \approx 2.7495$$

- $\delta = \mathbb{E}(|X-m|^3) = \frac{1}{10}|0-1.8|^3 + \frac{8}{10}|1-1.8|^3 + \frac{1}{10}|10-1.8|^3 =$

$$= \frac{1}{10}\left(1.8^3 + \cancel{8\cdot 8.2^3} + 8\cdot 0.8^3 + 8.2^3\right) = 56.1296$$

$$\Rightarrow \boxed{\text{hiba} \leq \frac{C\delta}{\sqrt{n}\,\sigma^3} \approx \frac{0.4748\cdot \cancel{2.7495}\ 56.1296}{100\cdot 2.7495^3} \approx 0.0128 = 1.28\%}$$

~~[illegible]~~.

④ Legyen X_i az i-edik kérés kiszolgálásához szükséges idő ms-ban mérve, $i=1,2,\dots,n$, ahol $n=10000$, és legyen $S_n = X_1 + X_2 + \dots + X_n$. A feladat a $\mathbb{P}(S_n \geq 20000)$ valószínűség nagy eltérés becslése.

Az X_i-k függetlenek és azonos eloszlásúak, közös eloszlásuk

k	0	1	10
$\mathbb{P}(X_i=k)$	$\frac{1}{10}$	$\frac{8}{10}$	$\frac{1}{10}$

explicitem ismert.

Nagy eltérés becslésre a Hoeffding-egyenlőtlenség és a Cramér-tétel is alkalmazható lenne, de X_i eloszlása nem nevezetes, így a rátafüggvény kiszámolása macerás lenne. Könnyebb a Hoeffding-egyenlőtlenség. Ehhez

- Korlátok: $0 =: a_i \leq X_i \leq \cancel{10}\, b_i := 10$, így
$$\sum_{i=1}^{n} (b_i - a_i)^2 = n\,(10-0)^2 = 10^4 \cdot 100 = 10^6$$
- várható érték: $\mathbb{E}X_i = \frac{1}{10}\cdot 0 + \frac{8}{10}\cdot 1 + \frac{1}{10}\cdot 10 = 1.8$, így
$$\mathbb{E}S_n = n\,\mathbb{E}X_1 = 10000 \cdot 1.8 = 18000$$

vagyis $\underline{t := 2000}$ választással

$$\boxed{\mathbb{P}(S_n \geq 20000)} = \mathbb{P}(S_n \geq \mathbb{E}S_n + t) \overset{\text{Hoeffding}}{\leq} \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right) =$$

$$= \exp\left(-\frac{2\cdot 2000^2}{10^6}\right) = e^{-8} \simeq 0.000335 = \boxed{0.0335\,\%}$$

5

<u>1. megoldás</u> Legyen $n=100000$ és legyen $i=1,2,\dots,n$-re

~~$X_i = \{1, h_n$ az i-edik üzenet hib~~

X_i: az i-edik csomag <u>sikeres</u> átküldéséhez szükséges próbálkozások száma. A feladat szövege szerint $X_i \sim Geom(p)$ ahol $p = 0.999$ a hibátlan átvitel valószínűsége. $\mathbb{E}X_i = \frac{1}{p} = 1.001\overline{001}\dots =: m$

Legyen $S_n = X_1 + \dots + X_n$ az összes üzenet száma, a kérdés így $\mathbb{P}(S_n > 100200)$ nagy eltérés becslése.

Az X_i-k függetlenek és azonos eloszlásúak, így a Cramér tétel szerint

$$\mathbb{P}(S_n > 100200) = \mathbb{P}\left(\frac{S_n}{n} > 1.002\right) = \mathbb{P}\left(\frac{S_n}{n} \in (a,b)\right) \lessapprox$$

$a := 1.002,\ b := \infty$ mivel $\boxed{a > m}$

$$\lessapprox e^{-nI(a)} = e^{-10^5 \cdot I(1.002)},$$ ahol I a $p = 0.999$ paraméterű geometriai eloszlás Cramér féle rátafüggvénye.

$$I(x) = (x-1)\ln\frac{x-1}{1-p} - x\ln x - \ln p = (x-1)\ln\frac{x-1}{0.001} - x\ln x - \ln 0.999$$

$$\Rightarrow I(1.002) = 0.002\ln\frac{0.002}{0.001} - 1.002\ln 1.002 - \ln 0.999 \approx 3.84796 \cdot 1e^{-4}$$

$$\Rightarrow nI(1.002) \approx 38.4796$$

$$\Rightarrow \boxed{\mathbb{P}(S_n > 100200) \lessapprox e^{-38.4796} \approx 1.94 \cdot 10^{-17}}$$

5) 2. megoldás Legyen $n = 100000$ és legyen $i = 1, 2, \dots, n$-re X_i az i-edik csomag sikeres átküldését megelőző sikertelen próbálkozások száma — vagyis a sikert megelőző kudarcok száma. A feladat szövege szerint

$X_i \sim PesszGeom(p)$, ahol $p = 0.999$ a siker-valószínűség.

$\mathbb{E}X_i = \frac{1}{p} - 1 = 0.001\overline{001}\dots \approx 0.001.$

Legyen $S_n = X_1 + X_2 + \dots + X_n$ az összes hibás üzenet száma (mire az $n = 100000$ ~~hibátlan üze~~ csomag hibátlanul átmegy),

a kérdés így $\mathbb{P}(S_n > 200)$ nagy eltérés becslése.

Az X_i-k függetlenek és azonos eloszlásúak, így a Cramér tétel szerint

$$\mathbb{P}(S_n > 200) = \mathbb{P}\left(\frac{S_n}{n} > 0.002\right) = \mathbb{P}\left(\frac{S_n}{n} \in (a, b)\right) \underset{\text{mivel } a > m}{\lessapprox} e^{-n I(a)} =$$

$a := 0.002, \; b := \infty$

$$= e^{-10^5 \cdot I(0.002)}$$, ahol I a $PesszGeom(p = 0.999)$ eloszlás Cramér féle rátafüggvénye: $I(x) = x \ln \frac{x}{1-p} - (1+x)\ln(1+x) - \ln p$

$$\Rightarrow I(0.002) = 0.002 \ln \frac{0.002}{0.001} + 1.002 \ln(1.002) - \ln 0.999 \approx 3.84796 \cdot 10^{-4}$$

$$\Rightarrow n\, I(0.002) \approx 38.4796$$

$$\Rightarrow \boxed{\mathbb{P}(S_n > 200) \lessapprox e^{-38.4796} \approx 1.94 \cdot 10^{-17}}$$

(5) <u>3. megoldás</u> A kérdés úgy is feltehető, hogy $\boxed{n := 100200}$ próbálkozásból vajon sikerül-e legalább 100000 hibátlan küldést teljesíteni, vagy a hibák száma 200-nál nagyobb lesz.

Ezért legyen $i = 1, 2, \ldots, n$-re

$$X_i = \begin{cases} 1, & \text{ha az } i\text{-edik üzenet hibásan érkezik} \\ 0, & \text{ha nem} \end{cases}$$

Így $X_i \sim B(p)$ ahol $p = 0.001$ a hiba-valószínűség.

(Ebből ~~$\mathbb{E}X_i$~~ $m := \mathbb{E}X_i = p = 0.001$)

~~a kérdés $\mathbb{P}$~~

Legyen $S_n = X_1 + \ldots + X_n$ a hibák száma, így a feladat $\mathbb{P}(S_n > 200)$ nagy eltérés becslése.

<u>3/a. megoldás:</u> $a_i := 0$, $b_i := 1$ választással $a_i \leq X_i \leq b_i$, tehát a <u>Hoeffding-egyenlőtlenségből</u> $\mathbb{E}S_n = np = 100.2$, $t := \cancel{100}\ 99.8$ választással

$$\mathbb{P}(S_n > 200) = \mathbb{P}(S_n > \mathbb{E}S_n + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right) = e^{-\frac{2t^2}{n}}$$

$$= e^{-\frac{2\cdot \cancel{100^2}\, 99.8^2}{100200}} = e^{-0.199} \approx 0.82 = 82\%$$ $\boxed{\text{NAGYON ROSSZ BECSLÉS,}}$

nagy eltérés becslésnek <u>használhatatlan</u>

<u>3/b. megoldás:</u> Az X_i-k függetlenek, azonos eloszlásúak, így a <u>Cramér-tétel</u> szerint

$$\mathbb{P}(S_n > 200) = \mathbb{P}\left(\frac{S_n}{n} > \cancel{0.002}\ 1/501\right) = \mathbb{P}\left(\frac{S_n}{n} \in (a, b)\right) \lessapprox e^{-nI(a)}$$

$a = \frac{200}{100200} \approx 0.001996$ ~~$a := \frac{1}{501}$~~ $\boxed{\cancel{a := 0.002}}$, $b := \infty$

$a > m$

ahol $I(x) = x \ln\frac{x}{p} + (1-x)\ln\frac{1-x}{1-p} \Rightarrow I(\cancel{0.002}\ 1/501) \approx \cancel{3.867\cdot 10}^{-4}\ 3.8402797$

$\Rightarrow nI(a) \approx 38.4796 \Rightarrow \boxed{\mathbb{P}(S_n > 200) \lessapprox e^{-38.4796} \approx 1.94\cdot 10^{-17}}$