Nieoczekiwane konsekwencje zmiany czasu

czyli o tym, jak zmiana czasu zmienia niezmienniki

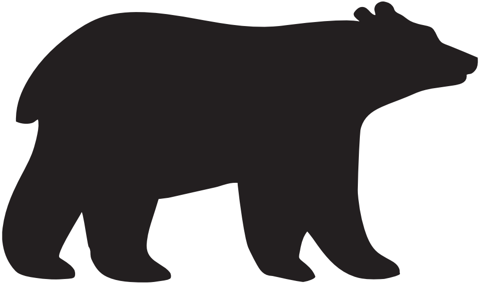
Andrzej Komisarski

andkom@math.uni.lodz.pl

WYDZIAŁ MATEMATYKI i INFORMATYKI
Uniwersytet Łódzki

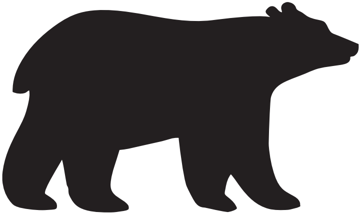
LXI Szkoła Matematyki Poglądowej
24 lutego 2020 r.
Wola Ducka

Motywacja: B.B., LIX SMP, luty 2019, poniedziałek, 9:00-9:45



Zadanie 11, zawody I stopnia L Olimpiady Matematycznej

W urnie znajdują się dwie kule: biała i czarna.
Ponadto mamy do dyspozycji 50 kul białych i 50 czarnych.
Wykonujemy 50 razy następującą czynność:
losujemy z urny kulę, a następnie wrzucamy ją z powrotem do urny oraz dokładamy jedną kulę tego samego koloru, co wylosowana kula.
Po zakończeniu tych czynności mamy więc w urnie 52 kule.
Jaka liczba kul białych znajdujących się w urnie jest najbardziej prawdopodobna?



Motywacja motywatora (B.B., LIX SMP, luty 2019, poniedziałek, 9:00-9:45)

Zadanie 11, zawody I stopnia L Olimpiady Matematycznej

W urnie znajdują się dwie kule: biała i czarna.
Ponadto mamy do dyspozycji 50 kul białych i 50 czarnych.
Wykonujemy 50 razy następującą czynność:
losujemy z urny kulę, a następnie wrzucamy ją z powrotem do urny oraz dokładamy jedną kulę tego samego koloru, co wylosowana kula.
Po zakończeniu tych czynności mamy więc w urnie 52 kule.
Jaka liczba kul białych znajdujących się w urnie jest najbardziej prawdopodobna?

Bardziej interesujące pytanie

Załóżmy, że mamy nieskończenie wiele dodatkowych kul i podaną w treści zadania czynność wykonujemy n razy. Jaki jest graniczny rozkład proporcji kul białych w urnie, gdy $n \to \infty$?

Zadanie 11, zawody I stopnia L Olimpiady Matematycznej

W urnie znajdują się dwie kule: biała i czarna.
Ponadto mamy do dyspozycji 50 kul białych i 50 czarnych.
Wykonujemy 50 razy następującą czynność:
losujemy z urny kulę, a następnie wrzucamy ją z powrotem do urny oraz dokładamy jedną kulę tego samego koloru, co wylosowana kula.
Po zakończeniu tych czynności mamy więc w urnie 52 kule.
Jaka liczba kul białych znajdujących się w urnie jest najbardziej prawdopodobna?

Bardziej interesujące pytanie

Załóżmy, że mamy nieskończenie wiele dodatkowych kul i podaną w treści zadania czynność wykonujemy n razy. Jaki jest graniczny rozkład proporcji kul białych w urnie, gdy $n \to \infty$?

Jak zmieni się odpowiedź, gdy zamiast z jedną kulą białą i jedną czarną zaczniemy z b kulami białymi i c czarnymi?

Motywacja motywatora (B.B., LIX SMP, luty 2019, poniedziałek, 9:00-9:45)

Zadanie 11, zawody I stopnia L Olimpiady Matematycznej

W urnie znajdują się dwie kule: biała i czarna.
Ponadto mamy do dyspozycji 50 kul białych i 50 czarnych.
Wykonujemy 50 razy następującą czynność:
losujemy z urny kulę, a następnie wrzucamy ją z powrotem do urny oraz dokładamy jedną kulę tego samego koloru, co wylosowana kula.
Po zakończeniu tych czynności mamy więc w urnie 52 kule.
Jaka liczba kul białych znajdujących się w urnie jest najbardziej prawdopodobna?

Bardziej interesujące pytanie

Załóżmy, że mamy nieskończenie wiele dodatkowych kul i podaną w treści zadania czynność wykonujemy n razy. Jaki jest graniczny rozkład proporcji kul białych w urnie, gdy $n \to \infty$?

Jak zmieni się odpowiedź, gdy zamiast z jedną kulą białą i jedną czarną zaczniemy z b kulami białymi i c czarnymi? A jeśli kolorów będzie więcej?

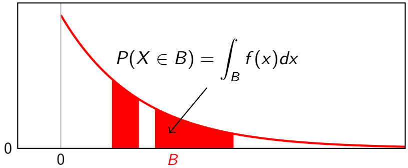
Schemat urnowy Pólya'i

Rozkład wykładniczy $\mathcal{E}(\beta)$

$\beta > 0$ – parametr

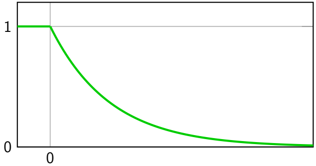
Gęstość

$$f(x) = \begin{cases} \beta \cdot e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$



Ogon rozkładu

$\overline{F}(x) = 1 - F(x) = P(X > x) = e^{-\beta x}$
dla $x \geq 0$



Wartość oczekiwana i wariancja

$EX = \frac{1}{\beta}$ $\quad$ $D^2X = \frac{1}{\beta^2}$

Skalowanie

Jeśli $X \sim \mathcal{E}(\beta)$ i $a > 0$, to $aX \sim \mathcal{E}(\beta/a)$.

Własność braku pamięci

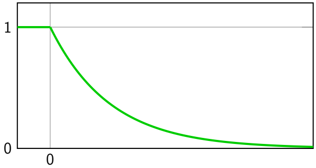
Jeśli $X \sim \mathcal{E}(\beta)$ i $s, t > 0$, to $P(X > s + t | X > s) = P(X > t)$.

Rozkład wykładniczy $\mathcal{E}(\beta)$

$\beta > 0$ – parametr

Ogon rozkładu

$\overline{F}(x) = 1 - F(x) = P(X > x) = e^{-\beta x}$
dla $x \geq 0$



Własność braku pamięci

Jeśli $X \sim \mathcal{E}(\beta)$ i $s, t > 0$, to $P(X > s + t | X > s) = P(X > t)$.

Minimum niezależnych zmiennych losowych o rozkładzie wykładniczym

Jeśli $X_1, \ldots, X_n$ są niezależne i $X_i \sim \mathcal{E}(\beta_i)$ dla $i = 1, \ldots, n$, to $\min(X_1, \ldots, X_n) \sim \mathcal{E}(\beta_1 + \cdots + \beta_n)$.

Dowód:
Dla $x \geq 0$ mamy
$P(\min(X_1, \ldots, X_n) > x) = P(X_1 > x, \ldots, X_n > x) =$
$= P(X_1 > x) \cdot \cdots \cdot P(X_n > x) = e^{-\beta_1 x} \cdot \cdots \cdot e^{-\beta_n x} = e^{-(\beta_1 + \cdots + \beta_n)x}$

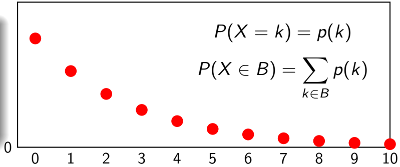
Rozkład geometryczny $Geom(p)$

$p \in (0, 1]$ – parametr

Jeśli wykonujemy nieskończony ciąg prób Bernoulliego (z prawdopodobieństwem sukcesu p), to liczba porażek przed pierwszym sukcesem ma rozkład $Geom(p)$.

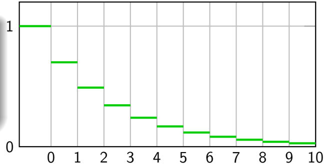
Funkcja prawdopodobieństwa

$$p(k) = \begin{cases} (1-p)^k \cdot p & k = 0, 1, \ldots \\ 0 & \text{w p.p.} \end{cases}$$



Ogon rozkładu

$$\overline{F}(k) = 1 - F(k) = P(X > k) = (1-p)^{k+1}$$
$$P(X \geq k) = (1-p)^k \text{ dla } k = 0, 1, \ldots$$



Wartość oczekiwana i wariancja

$$EX = \frac{1-p}{p} = \frac{1}{p} - 1 \qquad D^2X = \frac{1-p}{p^2}$$

Rozkład geometryczny $Geom(p)$

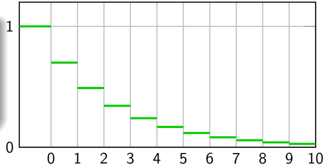
$p \in (0, 1]$ – parametr

Jeśli wykonujemy nieskończony ciąg prób Bernoulliego (z prawdopodobieństwem sukcesu p), to liczba porażek przed pierwszym sukcesem ma rozkład $Geom(p)$.

Ogon rozkładu

$\overline{F}(k) = 1 - F(k) = P(X > k) = (1-p)^{k+1}$

$P(X \geq k) = (1-p)^k$ dla $k = 0, 1, \ldots$



Własność braku pamięci

Jeśli $X \sim Geom(p)$, $i, j \in \{0, 1, \ldots\}$, to $P(X \geq i + j | X \geq i) = P(X \geq j)$.

Jeden ze związków rozkładu geometrycznego z rozkładem wykładniczym

Jeśli $X_1, X_2 \ldots$ oraz $m_1, m_2, \cdots > 0$ są takie, że $X_n \sim Geom(p_n)$, $m_n \to \infty$, $p_n \to 0$ oraz $m_n \cdot p_n \to \beta$, to dla dużych n zmienna losowa $\frac{X_n}{m_n}$ ma rozkład bliski $\mathcal{E}(\beta)$.

Dowód: $P(\frac{X_n}{m_n} > x) = P(X_n > \lfloor m_n x \rfloor) = (1 - p_n)^{\lfloor m_n x \rfloor + 1} \to e^{-\beta x}$.

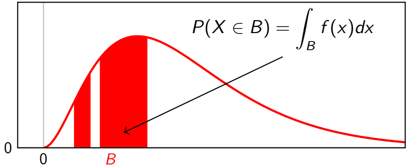
Rozkład gamma $\Gamma(\alpha, \beta)$

$\alpha, \beta > 0$ – parametry

Rozkład wykładniczy, to szczególny przypadek rozkładu gamma $\mathcal{E}(\beta) = \Gamma(1, \beta)$

Gęstość

$$f(x) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} \cdot e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$



Wartość ocz. i wariancja

$EX = \frac{\alpha}{\beta}$ $D^2X = \frac{\alpha}{\beta^2}$

Skalowanie

Jeśli $X \sim \Gamma(\alpha, \beta)$, $a > 0$, to $aX \sim \Gamma(\alpha, \beta/a)$.

Suma niezależnych zmiennych losowych o rozkładzie gamma

Jeśli $X_1, \ldots, X_n$ są niezależne i $X_i \sim \Gamma(\alpha_i, \beta)$ dla $i = 1, \ldots, n$, to $X_1 + \cdots + X_n \sim \Gamma(\alpha_1 + \cdots + \alpha_n, \beta)$.

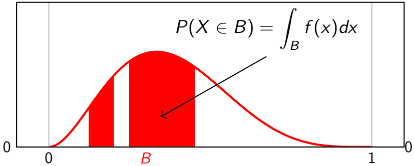
W szczególności, jeśli $X_1, \ldots, X_n \sim \mathcal{E}(\beta)$ są niezależne, to $X_1 + \cdots + X_n \sim \Gamma(n, \beta)$.

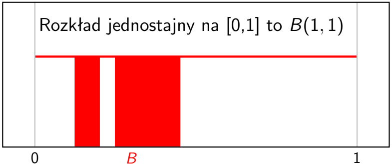
Rozkład beta $B(\alpha_1, \alpha_2)$

$\alpha_1, \alpha_2 > 0$ – parametry
Jeśli $X_1 \sim \Gamma(\alpha_1, \beta)$, $X_2 \sim \Gamma(\alpha_2, \beta)$ są niezależne, to $\frac{X_1}{X_1+X_2} \sim B(\alpha_1, \alpha_2)$ (oraz $\frac{X_2}{X_1+X_2} \sim B(\alpha_2, \alpha_1)$).

Gęstość

$$f(x) = \begin{cases} \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \cdot x^{\alpha_1-1}(1-x)^{\alpha_2-1} & x \in (0,1) \\ 0 & x \notin (0,1) \end{cases}$$





Wartość oczekiwana i wariancja

$EX = \frac{\alpha_1}{\alpha_1+\alpha_2}$ $\quad D^2X = \frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}$

Wracamy do zadania

Zadanie z przeróbkami

W urnie znajduje się b kul białych i c czarnych.
Wykonujemy n razy następującą czynność:
losujemy z urny kulę, a następnie wrzucamy ją z powrotem do urny oraz dokładamy jedną kulę tego samego koloru, co wylosowana kula.
Jaki jest graniczny rozkład proporcji kul białych w urnie, gdy $n \to \infty$?

Niech
B_k – liczba kul białych w urnie po k kolejkach
C_k – liczba kul czarnych w urnie po k kolejkach

$(B_k)_{k=0,1,\ldots}$, $(C_k)_{k=0,1,\ldots}$ to procesy stochastyczne (z czasem dyskretnym)

Wracamy do zadania

Zadanie z przeróbkami

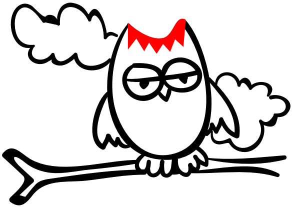
W urnie znajduje się b kul białych i c czarnych.
Wykonujemy n razy następującą czynność:
losujemy z urny kulę, a następnie wrzucamy ją z powrotem do urny oraz dokładamy jedną kulę tego samego koloru, co wylosowana kula.
Jaki jest graniczny rozkład proporcji kul białych w urnie, gdy $n \to \infty$?

Niech
B_k – liczba kul białych w urnie po k kolejkach
C_k – liczba kul czarnych w urnie po k kolejkach

$(B_k)_{k=0,1,\ldots}$, $(C_k)_{k=0,1,\ldots}$ to procesy stochastyczne (z czasem dyskretnym)

Niezmiennik: $B_k + C_k - k = b + c$

B_k i C_k nie są niezależne



Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots,$ lecz w chwilach losowych!

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots,$ lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\quad \beta > 0$

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots$, lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\quad \beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots$, lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\quad \beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Trzecie losowanie wykonamy w chwili T_3, przy czym po drugim losowaniu odczekamy losowy czas $T_3 - T_2 \sim \mathcal{E}((b+c+2)\beta)$, niezależny od T_1, T_2.

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots,$ lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\quad \beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Trzecie losowanie wykonamy w chwili T_3, przy czym po drugim losowaniu odczekamy losowy czas $T_3 - T_2 \sim \mathcal{E}((b+c+2)\beta)$, niezależny od T_1, T_2.

...

k-te losowanie wykonamy w chwili T_k, przy czym po $k-1$-szym losowaniu odczekamy losowy czas $T_k - T_{k-1} \sim \mathcal{E}((b+c+k-1)\beta)$, niezależny od $T_1, \ldots, T_{k-1}$.

...

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \dots,$ lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Trzecie losowanie wykonamy w chwili T_3, przy czym po drugim losowaniu odczekamy losowy czas $T_3 - T_2 \sim \mathcal{E}((b+c+2)\beta)$, niezależny od T_1, T_2.

...

k-te losowanie wykonamy w chwili T_k, przy czym po $k-1$-szym losowaniu odczekamy losowy czas $T_k - T_{k-1} \sim \mathcal{E}((b+c+k-1)\beta)$, niezależny od $T_1, \dots, T_{k-1}$.

...

Niech N_t będzie liczbą losowań do chwili $t \geq 0$

Co się zmieniło?

Było		Jest
$(B_k)_{k=0,1,\dots}$	liczba kul białych	$(B_{N_t})_{t\geq 0}$
$(C_k)_{k=0,1,\dots}$	liczba kul czarnych	$(C_{N_t})_{t\geq 0}$
czas dyskretny $k = 0, 1, \dots$	czas	czas ciągły $t \geq 0$
$B_k + C_k - k$	fajny niezmiennik	już go nie ma

I na co nam ta cała zmiana?

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots,$ lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Trzecie losowanie wykonamy w chwili T_3, przy czym po drugim losowaniu odczekamy losowy czas $T_3 - T_2 \sim \mathcal{E}((b+c+2)\beta)$, niezależny od T_1, T_2.

...

k-te losowanie wykonamy w chwili T_k, przy czym po $k-1$-szym losowaniu odczekamy losowy czas $T_k - T_{k-1} \sim \mathcal{E}((b+c+k-1)\beta)$, niezależny od $T_1, \ldots, T_{k-1}$.

...

Niech N_t będzie liczbą losowań do chwili $t \geq 0$

Co się zmieniło?

Było		Jest
$(B_k)_{k=0,1,\ldots}$	liczba kul białych	$(B_{N_t})_{t\geq 0}$
$(C_k)_{k=0,1,\ldots}$	liczba kul czarnych	$(C_{N_t})_{t\geq 0}$
czas dyskretny $k = 0, 1, \ldots$	czas	czas ciągły $t \geq 0$
$B_k + C_k - k$	fajny niezmiennik	już go nie ma

I na co nam ta cała zmiana?

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots,$ lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Trzecie losowanie wykonamy w chwili T_3, przy czym po drugim losowaniu odczekamy losowy czas $T_3 - T_2 \sim \mathcal{E}((b+c+2)\beta)$, niezależny od T_1, T_2.

...

k-te losowanie wykonamy w chwili T_k, przy czym po $k-1$-szym losowaniu odczekamy losowy czas $T_k - T_{k-1} \sim \mathcal{E}((b+c+k-1)\beta)$, niezależny od $T_1, \ldots, T_{k-1}$.

...

Niech N_t będzie liczbą losowań do chwili $t \geq 0$

Co się zmieniło?

Było		Jest
$(B_k)_{k=0,1,\ldots}$	liczba kul białych	$(B_{N_t})_{t\geq 0}$
$(C_k)_{k=0,1,\ldots}$	liczba kul czarnych	$(C_{N_t})_{t\geq 0}$
czas dyskretny $k = 0, 1, \ldots$	czas	czas ciągły $t \geq 0$
$B_k + C_k - k$	fajny niezmiennik	już go nie ma

I na co nam ta cała zmiana?

Zmieniamy czas na losowy

Nie będziemy losować kul w chwilach $1, 2, \ldots$, lecz w chwilach losowych!

Pierwsze losowanie wykonamy w chwili $T_1 \sim \mathcal{E}((b+c)\beta)$. $\beta > 0$

Drugie losowanie wykonamy w chwili T_2, przy czym po pierwszym losowaniu odczekamy losowy czas $T_2 - T_1 \sim \mathcal{E}((b+c+1)\beta)$, niezależny od T_1.

Trzecie losowanie wykonamy w chwili T_3, przy czym po drugim losowaniu odczekamy losowy czas $T_3 - T_2 \sim \mathcal{E}((b+c+2)\beta)$, niezależny od T_1, T_2.

...

k-te losowanie wykonamy w chwili T_k, przy czym po $k-1$-szym losowaniu odczekamy losowy czas $T_k - T_{k-1} \sim \mathcal{E}((b+c+k-1)\beta)$, niezależny od $T_1, \ldots, T_{k-1}$.

...

Niech N_t będzie liczbą losowań do chwili $t \geq 0$

Co się zmieniło?

Było		Jest
$(B_k)_{k=0,1,\ldots}$	liczba kul białych	$(B_{N_t})_{t\geq 0}$
$(C_k)_{k=0,1,\ldots}$	liczba kul czarnych	$(C_{N_t})_{t\geq 0}$
czas dyskretny $k = 0, 1, \ldots$	czas	czas ciągły $t \geq 0$
$B_k + C_k - k$	fajny niezmiennik	już go nie ma

I na co nam ta cała zmiana? B_{N_t} i C_{N_t} są niezależne!!!

Tego nie było na slajdach

Teraz mówiłem o tym, że nasze losowanie kul z dokładaniem w losowych chwilach jest równoważne następującemu procesowi:
zamiast b kul białych i c czarnych mamy b bakterii białych i c czarnych. Bakterie rozmnażają się przez podział, przy czym dzielą się niezależnie i czasy oczekiwania każdej z nich (oraz ich potomków) na podział są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{E}(\beta)$.

Wynika stąd zarówno to, że dla każdego $t \geq 0$ *zmienne losowe* B_{N_t} *i* C_{N_t} *są niezależne, jak i to, że do pełnego zanalizowania problemu wystarczy rozważyć rozwój populacji pochodzącej od pojedynczej bakterii.*

Oprócz tego pokazałem dwie symulacje komputerowe procesu wzrostu liczby kul w urnie: z czasem oryginalnym i z czasem zmieniającym się losowo.

Jak się mnoży jedna kula?

Niech $X_1 \sim \mathcal{E}(\beta), X_2 \sim \mathcal{E}(2\beta), \ldots, X_k \sim \mathcal{E}(k\beta), \ldots$ – niezależne

$T_1 = X_1$
$T_2 = X_1 + X_2$
$\ldots$
$T_k = \sum_{i=1}^{k} X_i$
$\ldots$

$T_1, T_2, \ldots$ to momenty kolejnych podziałów

N_t – liczba podziałów do chwili $t \geq 0$

Jak się mnoży jedna kula?

Niech $X_1 \sim \mathcal{E}(\beta), X_2 \sim \mathcal{E}(2\beta), \dots, X_k \sim \mathcal{E}(k\beta), \dots$ – niezależne

$T_1 = X_1$
$T_2 = X_1 + X_2$
$\dots$
$T_k = \sum_{i=1}^{k} X_i$
$\dots$

$T_1, T_2, \dots$ to momenty kolejnych podziałów

N_t – liczba podziałów do chwili $t \geq 0$

Rozkład liczby podziałów

$N_t \sim Geom(e^{-\beta t})$

Jak się mnoży jedna kula?

Niech $X_1 \sim \mathcal{E}(\beta), X_2 \sim \mathcal{E}(2\beta), \ldots, X_k \sim \mathcal{E}(k\beta), \ldots$ – niezależne

$T_1 = X_1$
$T_2 = X_1 + X_2$
$\ldots$ $\qquad$ $T_1, T_2, \ldots$ to momenty kolejnych podziałów
$T_k = \sum_{i=1}^{k} X_i$
$\ldots$

N_t – liczba podziałów do chwili $t \geq 0$

Rozkład liczby podziałów

$N_t \sim Geom(e^{-\beta t})$

Dowód: $P(N_t \geq k) = P(T_k \leq t) = P(\sum_{i=1}^{k} X_i \leq t) =$

Lemat

$\sum_{i=1}^{k} X_i \sim \max(Y_1, \ldots, Y_k)$, gdzie $Y_1, \ldots, Y_k \sim \mathcal{E}(\beta)$ są niezależne.

$= P(\max(Y_1, \ldots, Y_k) \leq t) = P(Y_1 \leq t) \ldots P(Y_k \leq t) = (1 - e^{-\beta t})^k$

Jak się mnoży jedna kula?

Zatem po czasie t z jednej kuli robi się $N_t + 1$
$N_t \sim Geom(e^{-\beta t})$, a więc tla dużych t rozkład $\frac{N_t}{e^{\beta t}}$ i rozkład $\frac{N_t+1}{e^{\beta t}}$ są bliskie $\mathcal{E}(1)$

Tego nie było na slajdach

Teraz, opierając się na własnościach rozkładów gamma i beta, dokończyłem rachunek pokazujący, że szukany graniczny rozkład proporcji kul białych to rozkład beta $B(b, c)$.
Następnie, już po czasie, przedstawiłem dwa komentarze (tutaj w zmienionej formie, oryginalnie były w formie rysunków):

Pierwszy komentarz dotyczył dowolnej skończonej liczby kolorów (może ich być więcej niż dwa) i związku otrzymanego wyniku ze statystykami pozycyjnymi z próby z rozkładu jednostajnego na przedziale $[0, 1]$. W przykładzie użyłem trzech kolorów: na początku mamy b kul białych, c czerwonych i n niebieskich. Rozkład graniczny (gdy $t \to \infty$) wektora (proporcja liczby kul białych, proporcja liczby kul czarnych, proporcja liczby kul niebieskich) jest taki sam, jak rozkład wektora $(X_{(b)} - 0, X_{(b+c)} - X_{(b)}, 1 - X_{(b+c)})$, gdzie $X_{(1)}, X_{(2)}, \ldots, X_{(b+c+n-1)}$ to posortowane od najmniejszej do największej wartości $X_1, X_2, \ldots, X_{b+c+n-1}$ wylosowane niezależnie z przedziału $[0, 1]$ zgodnie z rozkładem jednostajnym.

Tego nie było na slajdach

Drugi komentarz dotyczył porównania dwóch eksperymentów statystycznych.

W pierwszym hodowaliśmy przez czas T pojedynczą bakterię, której czas do podziału miał rozkład wykładniczy $\mathcal{E}(\beta)$. Notowaliśmy czasy kolejnych podziałów i rysowaliśmy „drzewo genealogiczne" powstającej kolonii bakterii.

W drugim, zapalaliśmy żarówkę, której czas życia miał rozkład $\mathcal{E}(\beta)$ i czekaliśmy, aż się spali. Jeśli spaliła się przed czasem T, to notowaliśmy czas, braliśmy kolejną żarówkę i próbę powtarzaliśmy. Eksperyment przerywaliśmy, gdy któraś żarówka nie spaliła się przez czas T.

Okazuje się, że eksperymenty te są statystycznie równoważne, to znaczy da się wyniki dowolnego z nich przekształcić w taki sposób, że statystyk widząc przekształcone wyniki jednego eksperymentu i oryginalne wyniki drugiego z nich nie jest w stanie rozróżnić pochodzenia tych wyników (ze względu na ich identyczny rozkład prawdopodobieństwa). W naszym przypadku to przekształcenie polega na odpowiednim pocięciu (lub sklejeniu) drzewa genealogicznego kolonii bakterii i na kolejnej zmianie czasu – tym razem trzeba puścić czas „od tyłu".