

Zatem liczba bankierska pojawiająca się przy okazji zagadnienia de Méré ma wartość **25**.

Tymczasem, jak pisaliśmy we wstępie do tego opracowania, Méré był przekonany kierując się intuicją i bogatym doświadczeniem w grach hazardowych, że liczbą tą jest liczba **24**. W takim razie intuicja zawiodła go.

Rozkład Pólya i rozkład hipergeometryczny

Zacniemy od pewnego doświadczenia, które opiszemy metodami teorii prawdopodobieństwa. Pozwoli to nam doświadczenie to nazywać zjawiskiem losowym. Wyobraźmy sobie, że mamy urnę zawierającą n kul białych i m kul czarnych, a więc łącznie $s = n + m$ kul. Niech $k \geq 1$ i $r \in \mathbf{Z}$. Losujemy z urny kulę. Po odczytaniu jej koloru wrzucamy do urny $r + 1$ kul tego samego koloru, co wylosowana kula. Powtarzamy tę czynność k -razy.

Aby opisać to zjawisko losowe zdefiniujemy zmienną losową $\mathbf{X}$ taką, że

$$\mathbf{X}(\Omega) = \{0, 1, \dots, k\},$$

gdzie zdarzenie $\{\omega \in \Omega: \mathbf{X}(\omega) = j\}$ oznacza, że w k losowaniach kulę białą wylosowano j razy, $j = 0, 1, \dots, k$. Wyznamy rozkład zmiennej losowej $\mathbf{X}$.

Zdarzenie $\{\omega \in \Omega: \mathbf{X}(\omega) = j\}$ mówi, że w k losowaniach wylosowano dokładnie j kul białych i $k - j$ kul czarnych. Rozważmy na razie sytuację, kiedy mamy do czynienia ze zdarzeniem B_j , które oznacza, że w j pierwszych losowaniach wylosowano same kule białe. W takim razie w kolejnych losowaniach jakiegokolwiek. Wtedy

$$P(B_j) = \frac{n}{s} \cdot \frac{n+r}{s+r} \cdot \dots \cdot \frac{n+(j-1)r}{s+(j-1)r}.$$

Weźmy zdarzenie $B_j \cap C_{k-j}$, które , oznacza że w $k - j$ ostatnich losowaniach wylosowano tylko kule czarne natomiast w j pierwszych kule białe. Uwzględniając wzór na prawdopodobieństwo $P(B_j)$ dostaniemy

$$P(B_j \cap C_{k-j}) = \frac{n}{s} \cdot \frac{n+r}{s+r} \cdot \dots \cdot \frac{n+(j-1)r}{s+(j-1)r} \cdot \frac{m}{s+jr} \cdot \frac{m+r}{s+(j+1)r} \cdot \dots \cdot \frac{m+(k-j-1)r}{s+(k-1)r}.$$

Zauważmy teraz, że kolejność pojawiających się kolorów wylosowanych kul nie jest istotna, bowiem skutkuje tylko zmianą kolejności czynników w liczniku powyższego wzoru. Poniważ liczba wszystkich rozmieszczeń j kul białych w ciągu długości k jest równa $\binom{k}{j}$, więc

$$p_j = P(\{\omega \in \Omega: \mathbf{X}(\omega) = j\}) = \binom{k}{j} P(B_j \cap C_{k-j}),$$

co daje rozkład $(p_j, j = 0, 1, \dots, k)$ zmiennej losowej $\mathbf{X}$

$$p_j = \binom{k}{j} \frac{n}{s} \cdot \frac{n+r}{s+r} \cdot \dots \cdot \frac{n+(j-1)r}{s+(j-1)r} \cdot \frac{m}{s+jr} \cdot \frac{m+r}{s+(j+1)r} \cdot \dots \cdot \frac{m+(k-j-1)r}{s+(k-1)r}.$$

Rozkład ten w literaturze nazywany jest *schematem Pólya* (patrz [4]). Powyższy wzór warto zapisać inaczej wykorzystując następujące podstawienia:

$$p = \frac{n}{s}, \quad q = \frac{m}{s}, \quad \alpha = \frac{r}{s}.$$

Zauważmy, że wtedy p oznacza prawdopodobieństwo wylosowania w pierwszym losowaniu kuli białej, q – kuli czarnej oraz $p + q = 1$.

Po zamianie zmiennych i uproszczeniach dostaniemy

$$p_j = \binom{k}{j} \frac{p(p+\alpha) \cdot \dots \cdot (p+(j-1)\alpha) \cdot q(q+\alpha) \cdot \dots \cdot (q+(s-j-1)\alpha)}{1 \cdot (1+\alpha)(1+2\alpha) \cdot \dots \cdot (1+(k-1)\alpha)}.$$

O rozkładzie (p_j) zmiennej losowej $\mathbf{X}$ będziemy mówili, że jest rozkładem Pólya o parametrach p, α, s, k i będziemy pisali $\mathbf{X} \in \mathcal{Po}(p, \alpha, s, k)$.

Można pokazać, że ([4])

Fakt 3.4.1 *Niech zmienna losowa $\mathbf{X} \in \mathcal{Po}(p, \alpha, s, k)$. Wtedy*

$$E\mathbf{X} = kp, \quad \text{var}(\mathbf{X}) = kpq \frac{1+k\alpha}{1+\alpha}.$$

Uwaga 3.4.2 *Rozkład Pólya pojawia się w zagadnieniach opisujących zjawisko rozprzestrzeniania się pandemii. Wtedy kolejny przypadek zachorowania zwiększa prawdopodobieństwo kolejnego zachorowania.*

Przy pewnych założeniach rozkład Pólya asymptotycznie zachowuje się jak rozkład dwumianowy. Zachodzi następujące twierdzenie

Twierdzenie 3.4.3 *Niech $\mathbf{X} \in \mathcal{Po}(p, \alpha, s, k)$, gdzie parametry: n, m, s, r mają tę własność, że:*

1.

$$\exists_{p \in (0,1)} \forall_{n,s} \frac{n}{s} = p,$$

2.

$$\lim_{s \rightarrow +\infty} \frac{r}{s} = 0.$$

Wtedy, przy $n, m \rightarrow +\infty$

$$\forall_j p_j \rightarrow \binom{k}{j} p^j q^{k-j}.$$

Ważnym przypadkiem schematu Pólya jest sytuacja kiedy $r = -1$. Oznacza to, że w zjawisku losowym, które opisaliśmy wyżej wylosowana kula nie jest zwracana do urny. W takim razie w tej sytuacji $0 \leq j \leq k \leq n$. Wtedy dla takich j, k po podstawieniu do wzoru na rozkład Pólya wartości $\alpha = -\frac{1}{s}$ dostaniemy nowy rozkład, który nazywamy *rozkładem hipergeometrycznym*. Wtedy po przekształceniach (patrz [4]) dla zmiennej losowej $\mathbf{X}$ o takim rozkładzie, będziemy pisali $\mathbf{X} \in \mathcal{H}(p, k, s)$, dostaniemy

$$p_j = P(\{\omega \in \Omega: \mathbf{X}(\omega) = j\}) = \frac{\binom{sp}{j} \binom{sq}{k-j}}{\binom{s}{k}}.$$

Fakt 3.4.2 Dla każdej zmiennej losowej $\mathbf{X} \in \mathcal{H}(p, k, s)$

$$E\mathbf{X} = kp, \quad \text{var}(\mathbf{X}) = \frac{s-k}{s-1}kpq.$$

Ostatnim z omawianych ważnym przykładem rozkładu dyskretnego jest:

Rozkład Poissona

Mówimy, że zmienna losowa $\mathbf{X}$ jest *typu Poissona*, jeśli

$$P(\{\omega \in \Omega : \mathbf{X}(\omega) = k\}) = e^{-\lambda} \frac{\lambda^k}{k!},$$

dla k całkowitych nieujemnych i dla pewnej liczby dodatniej λ . Dalej będziemy krótko pisali $\mathbf{X} \in \mathcal{P}(\lambda)$ (patrz Przykłady 2.3.1 i 3.3.5). Z Przykładu 3.3.5 wiemy, że $m = \lambda$ i $\sigma^2 = \lambda$.

Rolę rozkładu Poissona w probabilistyce podkreśla poniższe twierdzenie: