

Dekódování nejistoty: Diskrétní rozdělení pravděpodobnosti v byznysu

Jak proměnit náhodu v daty podloženou strategií.



Měření toho, co lze spočítat

Než můžeme modelovat budoucnost, musíme definovat naši proměnnou. V praxi pracujeme se dvěma světy: spojitým (měříme) a diskrétním (počítáme).

Spojitá situace

- Výška zákazníka
- Doba odezvy webu
- Objem kapalin



Diskrétní situace

- Počet konverzí
- Počet vadných kusů
- Denní počet zákazníků



Náš analytický arzenál se dnes zaměří na diskrétní jevy.

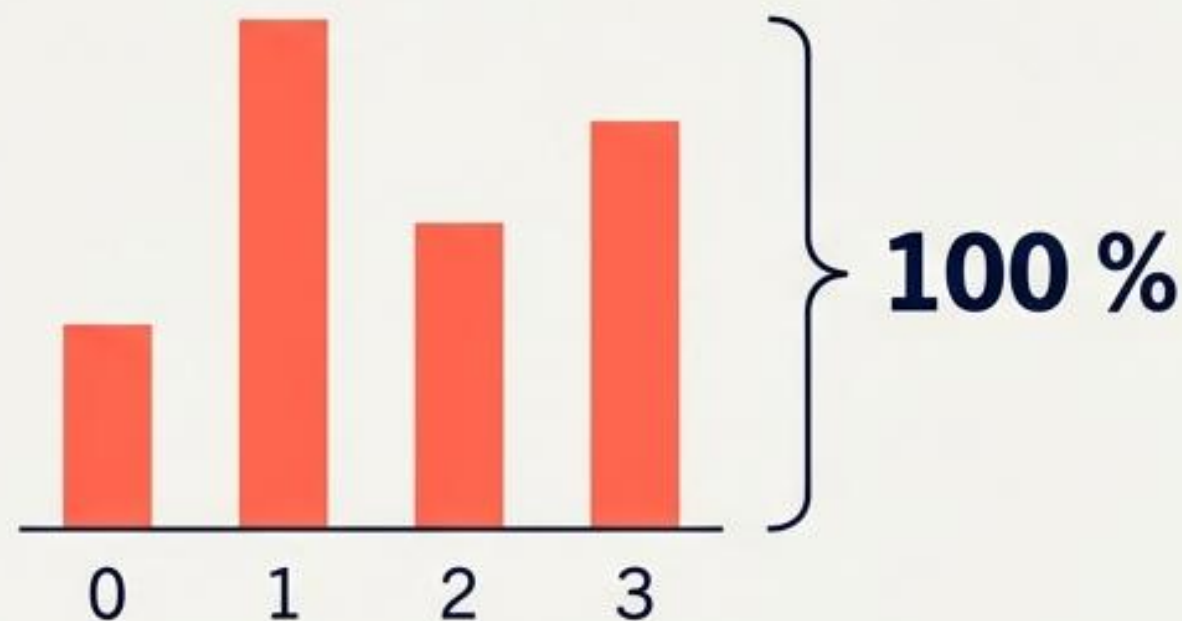
Základní kámen: Pravděpodobnostní funkce

$$P(X = x_i) = p_i$$

Pravděpodobnostní funkce přesně mapuje, s jakou šancí (p_i) nastane konkrétní výsledek (x_i).

$$\sum p(x) = 1$$

Základní pravidlo hry: Součet pravděpodobností všech možných scénářů musí vždy tvořit 100 %.



Nalezení středu: Očekávaná hodnota

Pokud bychom náš byznysový proces opakovali donekonečna, jaký bude průměrný výsledek? To je naše **Střední hodnota (Expected Value)**.

$$E(X) = \sum x_i \cdot P(X = x_i)$$



V byznysu představuje $E(X)$ očekávané příjmy, průměrný počet reklamací nebo cílový počet denních návštěv. Je to těžiště našeho modelu.

Kvantifikace rizika: Rozptyl a Směrodatná odchylka

Očekávaný výsledek není záruka. Rozptyl ukazuje, jak moc se realita může vychýlit od našeho plánu.

Rozptyl:

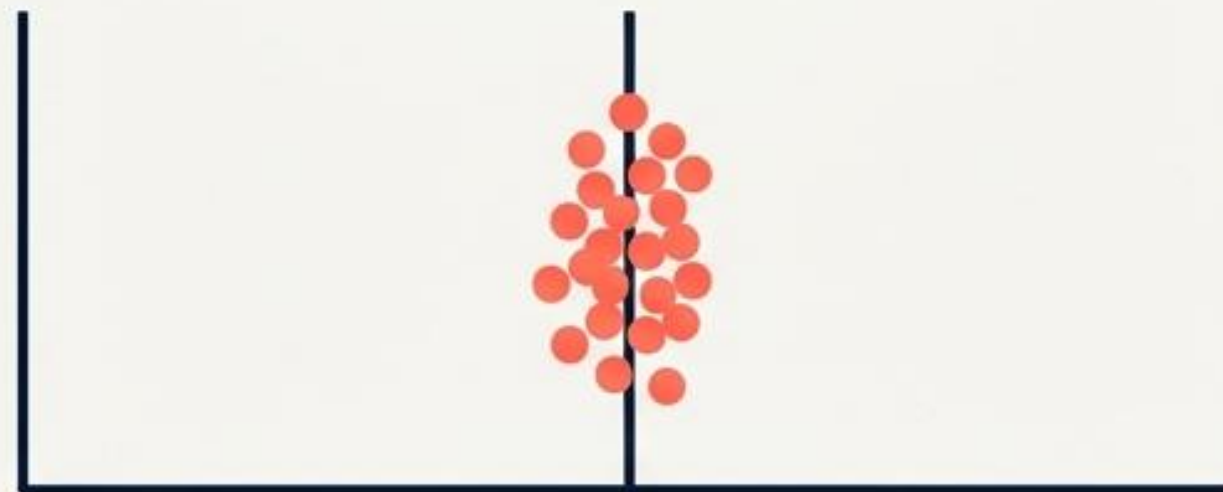
$$D(X) = \sum (x_i - E(X))^2 \cdot p_i$$

Směrodatná odchylka:

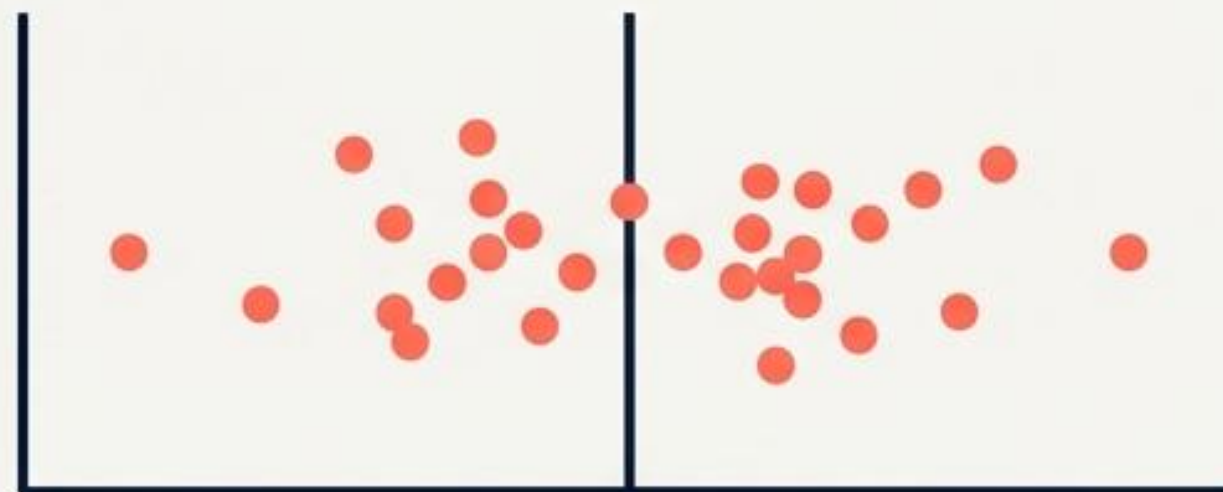
$$\sigma = \sqrt{D(X)}$$

Vyšší $D(X)$ znamená vyšší volatilitu trhu, nepředvídatelnější dodavatelský řetězec nebo vyšší riziko investice.

Nízké riziko (Malý rozptyl)



Vysoké riziko (Velký rozptyl)



Nástroj 1: Model Úspěch / Neúspěch

Binomické rozdělení $X \sim Bi(n, p)$

Modelujeme počet úspěchů v sérii n nezávislých pokusů, kde každý má přesně stejnou šanci na úspěch p .

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}, \quad k = 0, \dots, n, \quad n \geq 1, \quad p \in (0, 1)$$

Charakteristiky:

Střední hodnota: $E(X) = np$

Rozptyl: $D(X) = np(1 - p)$

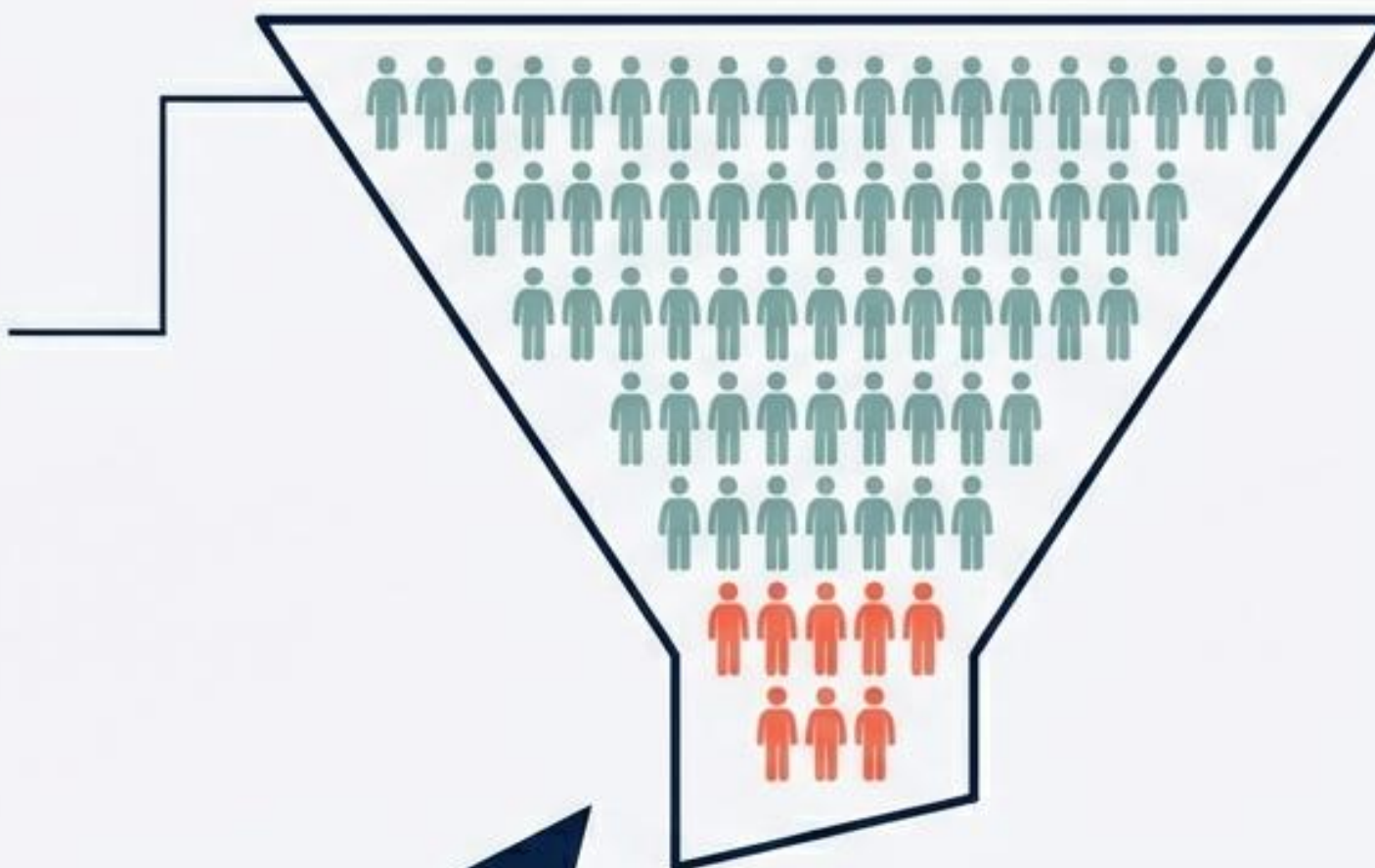
Binomické rozdělení v praxi

Scénář:

Marketingová konverzní kampaň.
Na váš e-shop přijde 100 návštěvníků
($n=100$). Dlouhodobý konverzní
poměr je 5 % ($p=0,05$).

Otázka:

Jaká je pravděpodobnost, že
získáme přesně 10 konverzí?



$$P(X=10) = 0,0167 \text{ (cca 1,67 \%)}$$

Excel: =BINOM.DIST(10; 100; 0,05; 0)

Nástroj 2: Model Kontrola kvality

Hypergeometrické rozdělení $X \sim \text{Hg}(N, M, n)$

Co když je náš vzorek omezený a každý výběr mění pravděpodobnost toho dalšího? (Výběr bez vracení ze závislých pokusů).

$$P(X=k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

Parametry:

- N = Celkový počet jednotek
- M = Počet jednotek se sledovanou vlastností
- n = Velikost našeho výběru



Hypergeometrické rozdělení v praxi

Scénář: Audit dodavatelského řetězce. Přijali jsme dávku 100 součástek ($N=100$). Víme, že 20 z nich je vadných ($M=20$). Provedeme náhodnou kontrolu 10 kusů ($n=10$).

Otázka: Jaká je pravděpodobnost, že najdeme přesně 3 zmetky?



$$P(X=3) = \frac{\binom{20}{3} \cdot \binom{80}{7}}{\binom{100}{10}} = 0,209 \text{ (20,9 \%)}$$

Excel: =HYPGEOM.DIST(3; 10; 20; 100; 0)

Nástroj 3: Model Příchozí události

Poissonovo rozdělení $X \sim \text{Po}(\lambda)$

Jak modelujeme izolované události, které nastávají náhodně a nezávisle během pevně daného časového úseku?

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda},$$

$$k = 0, 1, 2, \dots, \lambda > 0$$



Zajímavost modelu:

Střední hodnota i rozptyl jsou u tohoto rozdělení totožné.

$$E(X) = \lambda \text{ a } D(X) = \lambda$$

Poissonovo rozdělení v praxi

Scénář:

Dostupnost IT infrastruktury.
Vaše serverovna hlásí v průměru
2 výpadky za 24 hodin ($\lambda=2$).

Otázka:

Jaká je pravděpodobnost
bezchybného provozu
(přesně 0 výpadků)
během dnešního dne?



$$P(X=0) = \frac{2^0}{0!} \cdot e^{-2} = 0,135 \text{ (13,5 \%)}$$

Excel: =POISSON.DIST(0; 2; 0)

Nástroj 4: Model Stejná šance

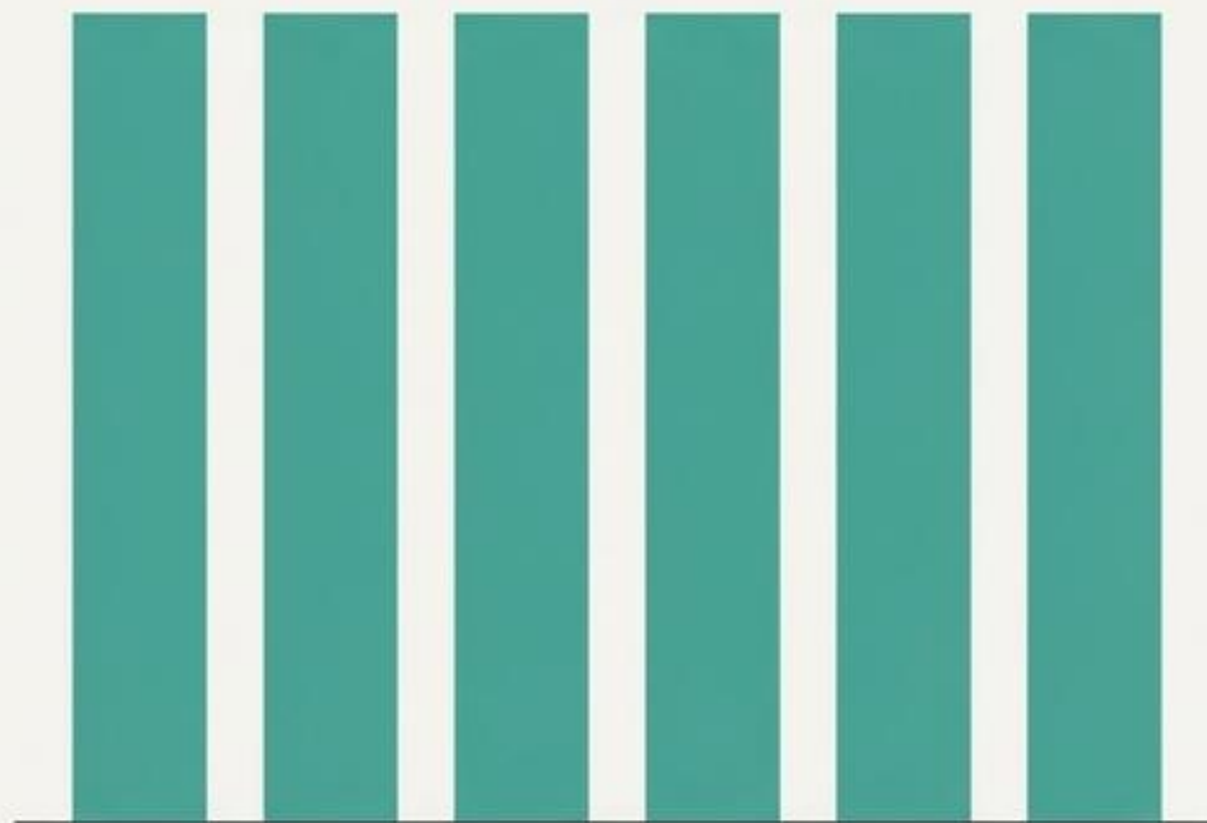
Diskrétní rovnoměrné rozdělení $X \sim \text{Ro}(n)$

Nejjednodušší model – nastává tehdy, když má každý z n možných výsledků naprosto stejnou šanci na úspěch.

$$P(X=x) = \frac{1}{n}$$

Využití v byznysu:

- Základní A/B testování (prvotní randomizace).
- Algoritmy pro rovnoměrné rozdělení zátěže (load balancing) mezi n zaměstnanců.



Shrnutí: Rozhodovací strom analytika

Který model použít k redukci vaší nejistoty?



Od chaosu ke strategii

Matematika není jen o vzorcích. Je to byznysová výhoda. Pochopením očekávané hodnoty $E(X)$ a rozptylu $D(X)$ a aplikací správného rozdělení transformujete nepředvídatelný chaos do kvantifikovatelného rizika.

**Nejistotu nelze eliminovat,
ale s těmito nástroji ji lze řídit.**

