

Pravděpodobnost pro humanitní vědce

Základy pravděpodobnosti a její aplikace

Obsah

1. Základy pravděpodobnosti – Nepodmíněná pravděpodobnost
2. Pravděpodobnost výběru vadných výrobků – Podmíněná pravděpodobnost
3. Bayesovská pravděpodobnost – Drogový pes
4. Narozeninový paradox
5. Rozšířené příklady z humanitních věd
6. Další aplikace pravděpodobnosti v humanitních vědách
7. Shrnutí klíčových bodů
8. Závěr a cvičení

1. Základy pravděpodobnosti

Nepodmíněná pravděpodobnost


Nepodmíněná pravděpodobnost je šance, že se něco stane, aniž by záviselo na jakékoli jiné události. Je to jedna z nejzákladnějších forem pravděpodobnosti a pomáhá nám chápat, jak často se určité výsledky očekávají.

Příklad: Hod mincí

- Když hodíme mincí, máme dvě možnosti: panna nebo orel.
- Nepodmíněná pravděpodobnost, že padne panna, je nezávislá na čemkoli jiném.
- Známe dvě možné výsledky, takže pravděpodobnost, že padne panna, je:

$$P(\text{panna}) = \frac{1}{2} = 50\%$$

Pravděpodobnostní strom

 Pravděpodobnostní strom { width="80%" }

Hod dvakrát za sebou

- Pokud hodíte mincí dvakrát, každé hodu je nezávislý na předchozím.
- Když chceme vědět, jaká je šance, že padne panna dvakrát za sebou, vynásobíme pravděpodobnosti pro jednotlivé hody:

$$P(\text{panna dvakrát}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 25\%$$

- Každý hod je nezávislý, takže šance na jakoukoliv kombinaci výsledků zůstává vždy 50 % pro jeden hod.

Vysvětlení v praxi

Nepodmíněná pravděpodobnost se hodí, když řešíme problémy, kde nezáleží na předchozích událostech. Například při sázení na výsledek hodu mincí víme, že šance jsou vždy stejné, ať už byla předtím panna nebo orel.

2. Pravděpodobnost výběru vadných výrobků

Podmíněná pravděpodobnost

Podmíněná pravděpodobnost počítá s tím, že určitá událost nastane pod podmínkou, že se již stala jiná událost. Jinými slovy, výsledek závisí na tom, co se už stalo.

Příklad: Vadné žárovky

- Máme krabici se 100 žárovkami, z nichž 9 je vadných.
- Chceme zjistit, jaká je pravděpodobnost, že vybereme dvě vadné žárovky za sebou.
- [Příklad zde](#)

1. **První výběr:** Pravděpodobnost, že vyberete vadnou žárovku je:

$$P(\text{vadná}) = \frac{9}{100}$$

2. **Druhý výběr:** Pokud jste už vybrali vadnou žárovku, zbývá jich 8 z 99. Proto je pravděpodobnost pro druhý výběr:

$$P(\text{vadná při druhém výběru}) = \frac{8}{99}$$

Výsledná pravděpodobnost

- Pravděpodobnost, že vyberete dvě vadné žárovky, spočítáme vynásobením jednotlivých kroků:

$$P(\text{dvě vadné}) = \frac{9}{100} \times \frac{8}{99} = 0,007 \quad (0,7$$

Vysvětlení v praxi

Podmíněná pravděpodobnost se často používá při rozhodování v reálném světě, kdy předchozí akce ovlivňují následující kroky. Například pokud jste již vybrali vadný výrobek, zvyšuje se pravděpodobnost, že další výběr bude ovlivněn tím, co jste už udělali.

3. Bayesovská pravděpodobnost – Drogový pes

Bayesova věta a podmíněná pravděpodobnost

Bayesova věta umožňuje zjistit pravděpodobnost nějaké události na základě dalších informací. Tím nám pomáhá propojit nové důkazy s tím, co už víme.

Příklad: Drogový pes

- Jak a kdy fungují AG testy
- Na letišti máme psa, který čichá zavazadla a hledá drogy. Pes má 98% úspěšnost při označování zavazadel s drogami, ale někdy se splete a štěká na zavazadla, kde drogy nejsou (8 % falešných poplachů).
- Na letišti je 5 % zavazadel, která skutečně obsahují drogy.

Výpočet Bayesovy pravděpodobnosti

Použijeme Bayesovu větu k výpočtu pravděpodobnosti, že v zavazadle jsou drogy, pokud pes štěká:

1. Představme si dva scénáře:

- Zavazadlo obsahuje drogy a pes štěká.
- Zavazadlo neobsahuje drogy a pes přesto štěká.

2. Výsledná pravděpodobnost:

$$P(\text{drogy} \mid \text{š tě kot}) = \frac{P(\text{š tě kot} \mid \text{drogy}) \times P(\text{drogy})}{P(\text{š tě kot})}$$

Výpočet nám dává šanci, že jsou v zavazadle drogy, když pes štěká, a to přibližně **39 %**.

Vysvětlení v praxi

Bayesova věta je velmi užitečná, když chceme propojit nové informace s tím, co už víme. V tomto případě víme, že pes obvykle správně štěká, když najde drogy, ale falešné poplachy snižují jistotu. Bayesova pravděpodobnost nám pomůže udělat lepší rozhodnutí.

4. Narozeninový paradox

Co je narozeninový paradox?

Narozeninový paradox se zdá neintuitivní, ale ukazuje, jak rychle roste pravděpodobnost, že dva lidé budou mít narozeniny ve stejný den, když je ve skupině více lidí.

[Článek zde](#)

Příklad: Skupina 23 lidí

- Jaká je pravděpodobnost, že dva lidé mají narozeniny ve stejný den ve skupině 23 lidí?
- Ačkoli se to zdá nepravděpodobné, pravděpodobnost je přibližně **50 %**!

Jak to funguje?

- Pravděpodobnost, že žádné dva osoby nemají narozeniny ve stejný den, se počítá postupně pro každou další osobu ve skupině.
- Jakmile je skupina dostatečně velká (v tomto případě 23 lidí), šance na shodu narozenin rychle vzrůstá.

Vysvětlení v praxi

Tento paradox je praktický zejména při analyzování situací, kde dochází k mnoha interakcím mezi prvky. V reálném světě to může být užitečné při plánování kapacit nebo při analýze dat z velkých skupin lidí.

5. Rozšířené příklady z humanitních věd

Pravděpodobnost v archeologii

Příklad: Nalezení artefaktů na archeologickém nalezišti

- Předpokládejme, že pravděpodobnost nalezení artefaktu v dané vrstvě je 0.3 (30%)
- Jaká je pravděpodobnost nalezení alespoň jednoho artefaktu při třech nezávislých výkopech?

$$P(\text{alespoň jeden}) = 1 - P(\text{ž á dný}) = 1 - (0.7)^3 \approx 0.657 \text{ nebo } 65.7\%$$

Podrobné vysvětlení:

1. **Komplementární pravděpodobnost:** Místo přímého výpočtu pravděpodobnosti nalezení artefaktu počítáme pravděpodobnost, že nenajdeme žádný, a odečteme ji od 1.
2. **Nezávislé pokusy:** Každý výkop je nezávislý, proto můžeme pravděpodobnosti násobit.
3. **Interpretace:** Přibližně v 65.7% případů najdeme alespoň jeden artefakt při třech výkopech. To je mnohem vyšší šance než při jediném výkopu (30%), což demonstruje význam opakovaných pokusů v archeologickém výzkumu.
4. **Aplikace v archeologii:** Tento typ výpočtu pomáhá archeologům plánovat rozsah vykopávek a odhadovat pravděpodobnost úspěchu expedice.

Pravděpodobnost v lingvistice

Příklad: Analýza autorství historických textů

- Předpokládejme, že známý autor používá určité slovo s frekvencí 0.01 (1%)
- V analyzovaném textu o 1000 slovech se toto slovo vyskytuje 20krát
- Je pravděpodobné, že text napsal tento autor?

Použijeme binomické rozdělení k výpočtu pravděpodobnosti:

$$P(X = 20) = \binom{1000}{20} (0.01)^{20} (0.99)^{980} \approx 4.3 \times 10^{-6}$$

Podrobné vysvětlení:

1. **Binomické rozdělení:** Toto rozdělení používáme, když máme fixní počet nezávislých pokusů (zde 1000 slov) a zajímá nás pravděpodobnost přesně k úspěchů (zde 20 výskytů slova).
2. **Interpretace výsledku:** Pravděpodobnost 4.3×10^{-6} je extrémně nízká, což naznačuje, že je velmi nepravděpodobné, že by text napsal známý autor.
3. **Stylometrie v praxi:** Tato metoda se používá v literární vědě pro analýzu autorství. Porovnáváme frekvence slov, frází nebo gramatických struktur s known corpus autora.
4. **Omezení:** Je důležité vzít v úvahu více faktorů, ne jen frekvenci jednoho slova. Komplexní stylometrické analýzy zohledňují mnoho různých lingvistických rysů.

6. Další aplikace pravděpodobnosti v humanitních vědách

Sociologie

- **Pravděpodobnostní vzorkování:** Použití náhodných vzorků k reprezentaci větší populace.
- **Analýza sociálních sítí:** Pravděpodobnost vzniku vazeb mezi jednotlivci.

Historie

- **Bayesovská analýza historických událostí:** Aktualizace pravděpodobnosti různých hypotéz na základě nových důkazů.

Psychologie

- **Teorie rozhodování:** Modelování pravděpodobnosti různých rozhodnutí v experimentálních podmínkách.

Antropologie

- **Modelování kulturní evoluce:** Pravděpodobnost šíření kulturních prvků v populaci.

Literární věda

- **Stylometrie:** Pravděpodobnostní analýza stylistických rysů pro určení autorství nebo datování textů.

7. Shrnutí klíčových bodů

1. Nepodmíněná pravděpodobnost:

- Základní forma pravděpodobnosti
- Nezávisí na jiných událostech

2. Podmíněná pravděpodobnost:

- Zohledňuje předchozí události
- Důležitá pro sekvenční rozhodování

3. Bayesova věta:

- Propojuje předchozí znalosti s novými důkazy
- Klíčová pro aktualizaci pravděpodobností

4. Narozeninový paradox:

- Ukazuje, jak rychle roste pravděpodobnost shody v větších skupinách
- Demonstruje neintuitivní aspekty pravděpodobnosti

5. Aplikace v humanitních vědách:

- Archeologie: Odhad pravděpodobnosti nálezů
- Lingvistika: Analýza autorství a frekvence slov
- Sociologie: Pravděpodobnostní vzorkování
- Historie: Bayesovská analýza událostí
- Psychologie: Modelování rozhodování
- Antropologie: Modelování kulturní evoluce
- Literární věda: Stylometrická analýza

8. Závěr a cvičení

Cvičení

1. Spočítejte pravděpodobnost, že při pěti hodech mincí padne orel pokaždé.
2. Navrhněte aplikaci pravděpodobnosti ve vašem oboru humanitních věd.

Děkuji za pozornost!