

Slump och sannolikhet

Random chance & Probability

Mathias Johansson Kristofer Söderström

2025-12-15

Innehållsförteckning

1	Vad är sannolikhet?	1
2	Exempel: Tärningskast	2
3	Slump i forskning	2
4	Stora talens lag	2
5	Normalfördelningen	2
6	Normalfördelningen - visualisering	2
7	Centrala gränsvärdessatsen	3
8	Konfidensintervall	3
9	Statistisk signifikans	3
10	Vanliga missförstånd	3
11	Samplingfel vs systematiska fel	3
12	Praktisk tillämpning	4
13	Referenser	4

1 Vad är sannolikhet?

Sannolikhet (probability) mäter hur troligt det är att en händelse inträffar.

$$P(A) = \frac{\text{antal gynnsamma utfall}}{\text{antal möjliga utfall}}$$

Värde mellan 0 (omöjligt) och 1 (säkert)

2 Exempel: Tärningeskast

Ett tärningeskast:

- $P(6) = 1/6 = 0.167 (16.7\%)$
- $P(\text{udda}) = 3/6 = 0.5 (50\%)$
- $P(\text{4}) = 4/6 = 0.667 (66.7\%)$

3 Slump i forskning

Varför är slumptäckning viktigt?

- **Urval:** Representativitet kräver slumptäckning urval
- **Experiment:** Randomisering minskar bias
- **Statistik:** Många metoder bygger på sannolikhetsteori

4 Stora talens lag

När antalet observationer ökar, närmrar sig det observerade medelvärdet det förväntade värdet.

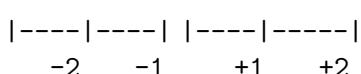
Exempel: Ju fler myntkast, desto närmare 50% krona

5 Normalfördelningen

Den klassiska "klockkurvan" (Gausskurvan)

- Medelvärdet i mitten
- Symmetrisk
- 68% inom ± 1 standardavvikelse
- 95% inom ± 2 standardavvikelse
- 99.7% inom ± 3 standardavvikelse

6 Normalfördelningen - visualisering



7 Centrala gränsvärdessatsen

Oavsett populationens fördelning blir urvalsmedelvärdet normalfördelade när urvalet är tillräckligt stort ($n > 30$).

Varför är detta viktigt?

Det låter oss använda normalfördelningen för att dra slutsatser om populationer.

8 Konfidensintervall

Ett intervall som med viss sannolikhet innehåller det sanna populationsvärdet.

95% konfidensintervall:

$$\bar{x} \pm 1.96 \times \frac{s}{\sqrt{n}}$$

“Vi är 95% säkra på att det sanna värdet ligger i detta intervall”

9 Statistisk signifikans

p-värde: Sannolikheten att få detta resultat (eller mer extremt) om nollhypotesen är sann.

- $p < 0.05$: “Statistiskt signifikant”
- $p < 0.01$: “Starkt signifikant”

OBS: Signifikans = praktisk betydelse!

10 Vanliga missförstånd

1. “Jag har tur/otur” → Slumpen har inget minne
2. “Det måste jämma ut sig” → Spelarens felslut (gambler’s fallacy)
3. “ $p = 0.05$ betyder 95% chans att vara sann” → Fel tolkning av p-värde

11 Samplingfel vs systematiska fel

– **Samplingfel (random error):** - Beror på slumpen i urvalet - Minskar med större urval - Kan skattas statistiskt

- **Systematiska fel (bias):**

- Beror på metodfel
- Minskar INTE med större urval
- Svårare att upptäcka

12 Praktisk tillämpning

Vid enkätundersökningar:

1. Hur stort urval behöver jag?
2. Vad är felmarginalen?
3. Kan jag generalisera resultaten?

Dessa frågor kräver förståelse för sannolikhet och slump.

13 Referenser

- Probability
- Normal distribution
- Central limit theorem
- Confidence interval
- Law of large numbers
- P-value
- 68–95–99.7 rule