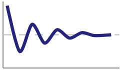
Probability and Random Variables

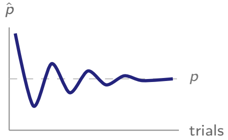
AP Statistics ▪ Unit 4

AP Statistics



What we will cover

1. Randomness and probability
2. Probability rules
3. Conditional and independence
4. Random variables
5. Combining variables
6. Binomial and geometric

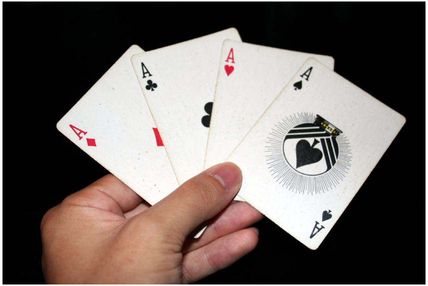


Randomness and probability

Probability 概率 is the long-run **relative frequency** 相对频率 of an outcome.

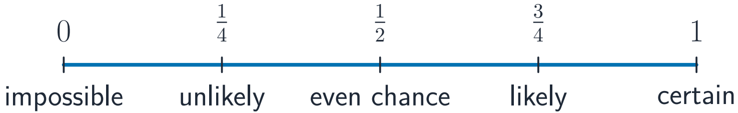
- short runs vary by **chance** 偶然
- **simulation** 模拟 estimates a hard probability

> **Law of large numbers** 大数定律: more trials $\Rightarrow$ $\hat{p}$ settles toward the true p.



playing cards

Randomness and probability



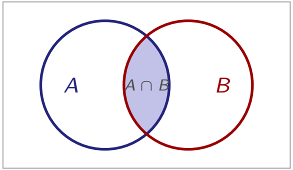
Probability rules

From the **sample space** 样本空间, every **event** 事件 obeys $0 \leq P \leq 1$, summing to 1.

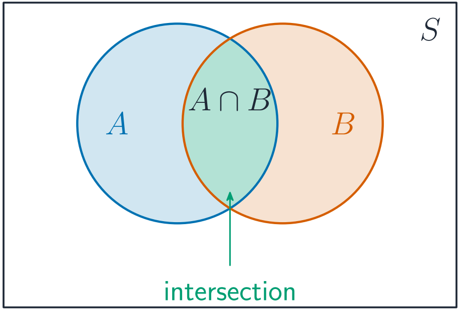
$$P(A^c) = 1 - P(A)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

> **Mutually exclusive** 互斥 events share nothing, so $P(A \cap B) = 0$ and the last term drops.



Probability rules



Conditional probability and independence

A **conditional probability** 条件概率 restricts to the times B happens:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

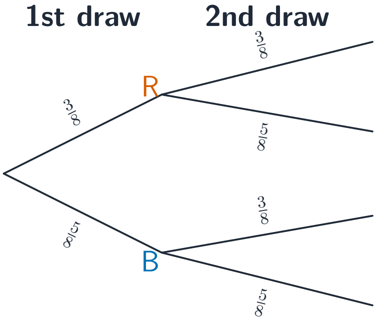
The **multiplication rule** 乘法法则:
$P(A \cap B) = P(A)\,P(B \mid A)$.

> **Independent** 独立: $P(A \mid B) = P(A)$, so $P(A \cap B) = P(A)P(B)$. Not the same as mutually exclusive!

	girl	boy
sport	24	36
none	16	24

$P(\text{girl} \mid \text{sport}) = \frac{24}{60}$

Conditional probability and independence



$$(\text{R}, \text{R}) = \frac{3}{8} \times \frac{3}{8} = \frac{9}{64}$$

$$(\text{R}, \text{B}) = \frac{15}{64}$$

$$(\text{B}, \text{R}) = \frac{15}{64}$$

$$(\text{B}, \text{B}) = \frac{25}{64}$$

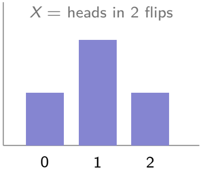
Random variables

A **random variable** 随机变量 X takes a number from a chance process (**discrete** 离散 or **continuous** 连续).

- its **distribution** 概率分布 lists each x with $P(x)$, summing to 1

$$\mu_X = \sum x_i P(x_i) \qquad \sigma_X^2 = \sum (x_i - \mu_X)^2 P(x_i)$$

> μ_X is the **expected value** 期望值 – the long-run average outcome.



Combining random variables

A **linear transformation** 线性变换:

$$\mu_{aX+b} = a\mu_X + b \qquad \sigma_{aX+b} = |a|\,\sigma_X$$

For a **sum or difference** 和或差 of **independent** 独立 variables, means combine and variances add – for *both*:

$$\mu_{X\pm Y} = \mu_X \pm \mu_Y \qquad \sigma^2_{X\pm Y} = \sigma^2_X + \sigma^2_Y$$

> Never subtract variances for a difference – variances always add.

Two packages

$\sigma_X = 0.3$, $\sigma_Y = 0.4$, independent.

$$\sigma_{X+Y} = \sqrt{0.3^2 + 0.4^2} = 0.5$$

The difference $X - Y$ has the *same* $\sigma = 0.5$.

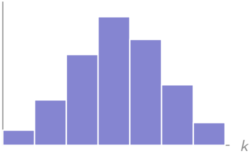
The binomial distribution

The **binomial setting** 二项分布情境 counts successes – remember **BINS**: **B**inary, **I**ndependent, fixed **N**umber, **S**ame p.

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

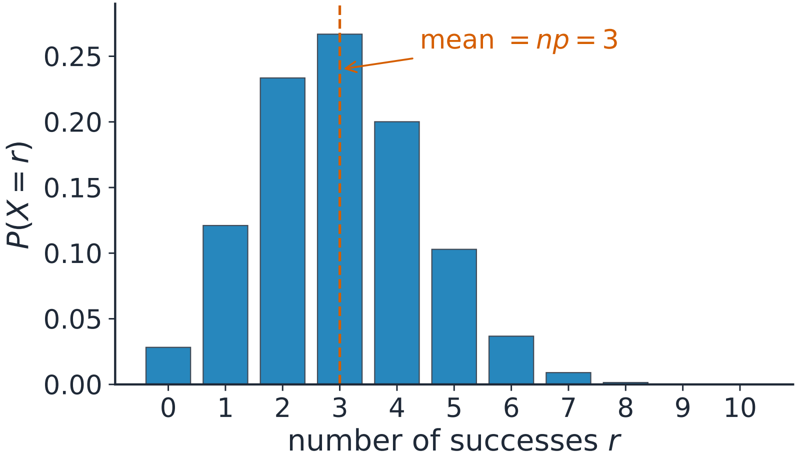
$$\mu_X = np \qquad \sigma_X = \sqrt{np(1 - p)}$$

> Shape: symmetric at $p = 0.5$, skewed otherwise; more bell-shaped as n grows.



Binomial: count of successes.

The binomial distribution



Worked example – binomial by hand

A free-throw shooter makes 80%. In $n = 10$ attempts, find $P(X = 8)$ and describe the distribution's centre and spread.

$$P(X = 8) = \binom{10}{8}(0.8)^8(0.2)^2$$

$$= 45(0.1678)(0.04) \approx 0.302$$

$\mu = np = 10(0.8) = 8$ makes

$\sigma = \sqrt{10(0.8)(0.2)} \approx 1.26$

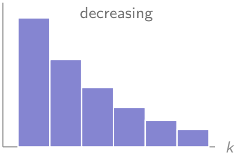
$\binom{10}{8} = 45$ counts which 8 shots dropped – forget it and you find one single ordering (0.0067), not the event. Check **BINS** first: without independence it is not binomial at all.

The geometric distribution

The **geometric setting** 几何分布情境 keeps going *until* the first success (no fixed n).

$$P(X = k) = (1-p)^{k-1}\,p \qquad \mu_X = \frac{1}{p}$$

> X counts trials until the first success; a binomial counts successes in a fixed number of trials.



Roll until a six: $\mu = 6$.

Unit 4 – key takeaways

1 **Probability**

long-run relative frequency; law of large numbers

2 **Rules**

addition for or; multiplication for and; conditional

3 **Random vars**

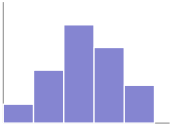
$\mu = \sum xP(x)$; variances add for independent

4 **Distributions**

binomial $\mu = np$; geometric $\mu = 1/p$

Check yourself

1. Write the complement rule for $P(A)$.
2. For independent X, Y, how do you combine their variances?
3. What does BINS stand for?
4. A binomial has $n = 20,\ p = 0.5$: what is μ_X?



Answers **1.** $P(A^c) = 1 - P(A)$ **2.** add them: $\sigma_X^2 + \sigma_Y^2$ **3.** Binary, Independent, Number, Same p **4.** $np = 10$