

Chapter 5

Probability

In this chapter, we look at basic properties of probability.

5.1 Probability Rules

We discuss law of large numbers, empirical probability and classical probability. We discuss statistical experiment, sample space and events. We define probability.

Exercise 5.1 (Probability Rules)

1. *Terminology.*

(a) *Summary.*

- *Law of large numbers:*
Repeatedly sampling, sample average (proportion) will approach and stay close to expected population average (probability).
- *Empirical probability:*
Probability event (E) in experiment approximated by
$$P(E) \approx \text{frequency of } E \div \text{number of trials of experiment.}$$
- *Classical probability:*
Probability event (E) in experiment, assuming equally likely outcomes,
$$P(E) = \text{number of ways } E \text{ occurs} \div \text{number outcomes in experiment.}$$
- *Sample space, S.*
List of all possible outcomes (or simple events) of experiment.
- *Event, E; simple event, e_i.*
Event is subset of all possible outcomes of experiment; simple event if subset consists of one outcome.
- *Probability experiment.*
Process which results in sample space where each outcome is assigned a chance of occurrence.

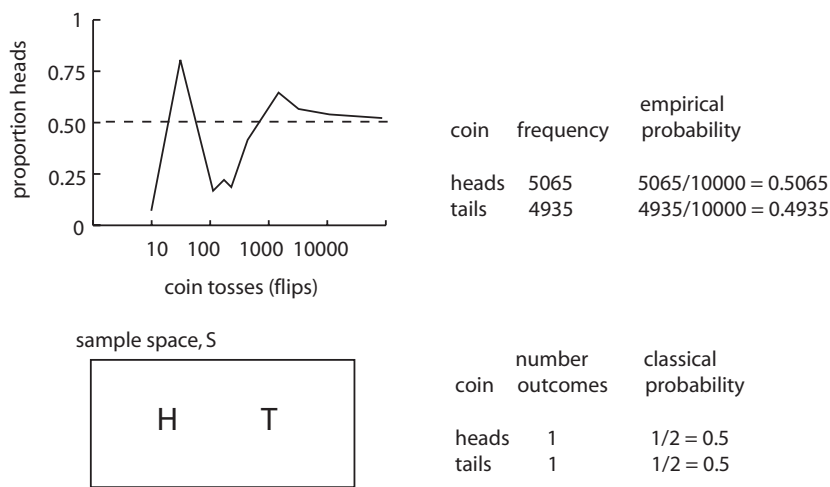
(b) *Coin tossing.*

Figure 5.1 (Terminology: coin tossing)

- i. **True / False.** As number of coin tosses increase, proportion of total tosses which are heads will approach and stay close to expected *probability* of tossing a head. This is an example of *law of large numbers*.
- ii. *Approximating probability with empirical approach*¹.
Since 5065 tosses of 10000 tosses are heads we approximate probability of tossing a head by $P(H) \approx \frac{5065}{10000} = \mathbf{0.4935} / \mathbf{0.5} / \mathbf{0.5065}$.
- iii. *Calculating probability with classical method*².
Since a coin can be tossed only as a head (H) or tail (T) and *assuming equally likely outcomes*, $P(H) = \frac{1}{2} = \mathbf{0.4935} / \mathbf{0.5} / \mathbf{0.5065}$.
- iv. *Classical method: experiment.*
Flipping a coin is a *probability experiment*. It is generally *unknown*, when flipping a coin, whether coin comes up heads (Hs) or tails (Ts). However, it is known there are only two possible outcomes (choose one) $\{\mathbf{H}, \mathbf{T}\} / \{\mathbf{H}, \mathbf{H}\} / \{\mathbf{T}, \mathbf{T}\}$.
- v. *Classical method: sample space*³.
Sample space for flipping a coin is $S = \{\mathbf{H}, \mathbf{T}\}$. Sample space is (choose one) **set** / **subset** / **element** of all possible outcomes.
- vi. *Classical method: event, simple event.*
Flipping a head $\{\mathbf{H}\}$ is an example of an *event*, E . An event is a (choose one) **set** / **subset** / **element** of all possible outcomes. Since only one outcome, $\{\mathbf{H}\}$, this event is a *simple event*, $E = e_1 = \{\mathbf{H}\}$

(c) *Die rolling.*¹Accuracy of empirical method depends on how well we can consistently flip the coin.²Accuracy of classical method depends on accuracy of equally likely outcomes assumption.³Displayed as a *Venn diagram* in figure.

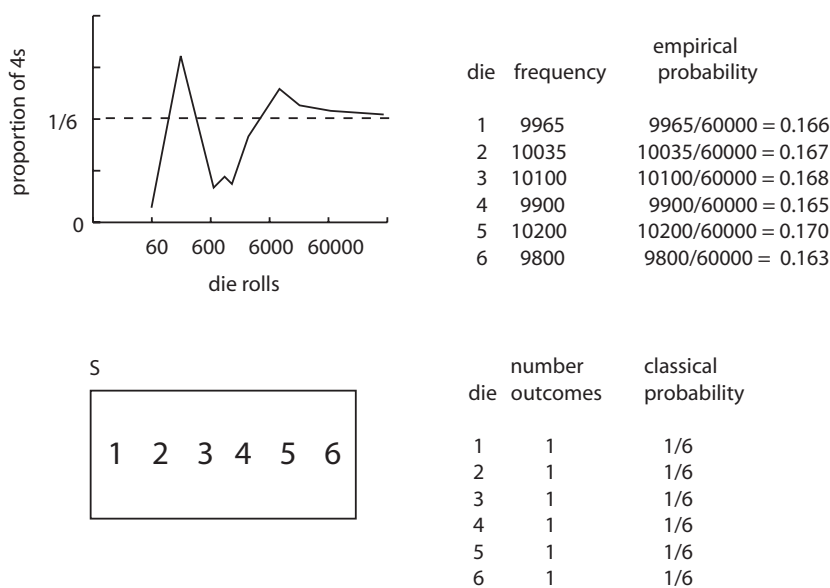


Figure 5.2 (Terminology: die rolling)

- i. **True / False.** As number of die rolls increase, proportion of total rolls which are 4s will approach and stay close to expected probability of rolling a 4. This is an example of *law of large numbers*.
 - ii. *Approximating probability with empirical approach.*
 Since 9900 tosses of 60000 rolls are 4s we approximate probability of rolling a 4 by $P(4) \approx \frac{9900}{60000} =$ (choose one) **0.163 / 0.164 / 0.165**.
 - iii. *Calculating probability with classical method.*
 Since a die can be rolled either 1, 2, 3, 4, 5 or 6 and assuming equally likely outcomes, $P(4) = \frac{1}{6} \approx$ (choose one) **0.163 / 0.165 / 0.167**.
 - iv. *Classical method: experiment.*
 Die rolling is *probability experiment*. Value to be rolled unknown, but six possible outcomes (choose one) **known / unknown**.
 - v. *Classical method: sample space.*
 Sample space is (choose one) **{1, 2, 3, 4, 5} / {1, 2, 3, 4, 5, 6}** .
 - vi. *Classical method: event, simple event.*
 Examples of events are
 (choose one or *more!*) **{1} / {1, 2} / {1, 2, 3, 6}** .
- (d) *Chance error, % chance error and law of large numbers: die rolling*
 Results of die rolling experiment given in table below.

no of rolls	no of 4s	expected 4s	difference	% difference
60	9	10	-1	$-\frac{1}{60} \times 100 \approx -1.7\%$
120	21	20	1	$\frac{1}{120} \times 100 \approx 0.8\%$
180	34	30	4	$\frac{4}{180} \times 100 \approx 2.2\%$
240	42	40	2	$\frac{2}{240} \times 100 \approx 0.8\%$
600	110	100	10	$\frac{10}{600} \times 100 \approx 1.7\%$
1200	234	200	34	$\frac{34}{1200} \times 100 \approx 2.8\%$
2400	390	400	-10	$-\frac{10}{2400} \times 100 \approx -0.4\%$

- i. After 60 rolls, **6 / 7 / 9** are 4s;
after 2400 rolls, **390 / 400 / 409** are 4s.
- ii. If fair, after 60 rolls, we *expect* $\frac{60}{6} = 4$ / **5 / 10** are 4s;
after 2400 rolls, we expect $\frac{2400}{6} = 390$ / **400 / 409** are 4s.
- iii. *Chance error* after 60 rolls $9 - 10 = -1$ / **0 / 1** are 4s;
after 2400 rolls, chance error $390 - 400 = -10$ / **-5 / 10**.
- iv. Chance error **small / large / same** for a small number of rolls but
smaller / same / larger for a large number of rolls.
- v. *Percentage chance error*, 60 rolls $\frac{9-10}{60} \times 100 = -1.7\%$ / **0% / 1.7%**;
after 2400 rolls, $\frac{390-400}{2400} \times 100 \approx -0.4\%$ / **-0.0% / 2.0%**.
- vi. Percentage of chance error **increases / decreases / remains same**
as number of rolls increases: *observed variability* around $\frac{1}{6}$ decreases.
- vii. So *observed* proportion of 4s **converges towards / diverges from**
the *expected* proportion of 4s, $\frac{1}{6}$, by the law of large numbers.
- viii. Law of large numbers **does / does not** describe the *expected variability*
of chance error around $\frac{1}{6}$, but is known to be a bell-shaped histogram.

2. Rules of probability.

(a) Rules.

- Probability of any event, E , must be between 0 and 1: $0 \leq P(E) \leq 1$.
- Sum of probability of all outcomes equals 1; for sample space $S = \{e_1, e_2, \dots, e_n\}$,

$$P(e_1) + P(e_2) + \dots + P(e_n) = 1.$$

(b) Flipping coin three times

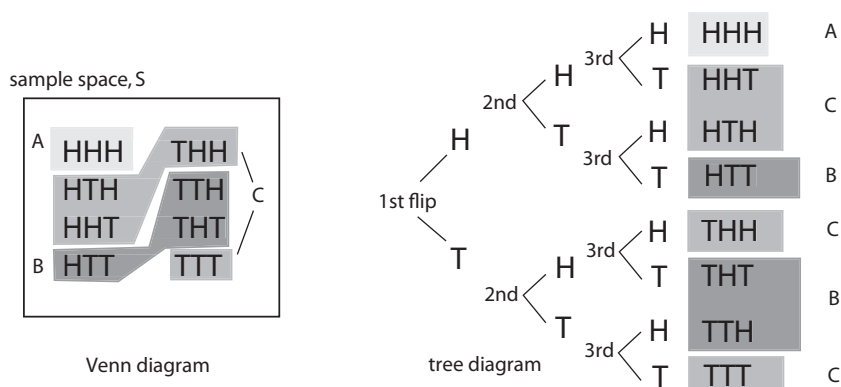


Figure 5.3 (Probability: flipping coin three times)

- i. Assuming each outcome equally likely,
 $P(A) = P(HHH) = (\text{circle one}) \frac{1}{8} / \frac{3}{8} / \frac{4}{8} / \frac{8}{1}$
- ii. Probability flipping *exactly one head*,
 $P(B) = P(HTT, THT, TTH) = (\text{circle one}) \frac{1}{8} / \frac{3}{8} / \frac{4}{8} / \frac{5}{8}$.
- iii. $P(C) = (\text{circle one}) \frac{1}{8} / \frac{4}{8} / \frac{6}{8} / \frac{8}{8}$.
- iv. Events A, B and C (choose one)
 - overlap or intersect**
 - do not overlap, are disjoint (mutually exclusive)**
 - from one another.
- v. Taken together, events A, B and C
 - cover or shade**
 - do not cover or do not completely shade**
 - entire sample space.
- vi. Rules of probabilities (choose one) **violated / obeyed**
 since $0 \leq P(E) \leq 1$, for $E = A, B, C$ and $P(A) + P(B) + P(C) = 1$.

(c) *Box of coins.*

Box contains ten coins, with following denominations and years: five cents (Cs) (1974, 1976, 1978, 1978 and 1980), three nickels (Ns) (1974, 1978 and 1980) and two dimes (Ds) (1978 and 1981).

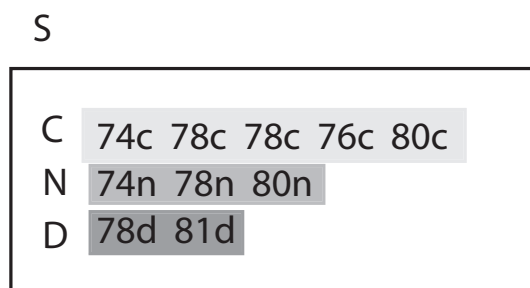


Figure 5.4 (Probability: choosing coins from a box)

- i. Assuming each outcome equally likely, chance choosing a *cent* (*penny*)
 $P(C) = P(74c, 78c, 78c, 76c, 80c) = (\text{circle one}) \frac{2}{10} / \frac{3}{10} / \frac{5}{10}$.
- ii. Probability of choosing nickel, $P(N) = (\text{circle one}) \frac{2}{10} / \frac{3}{10} / \frac{5}{10}$.
- iii. Probability of choosing dime, $P(D) = (\text{circle one}) \frac{2}{10} / \frac{3}{10} / \frac{5}{10}$.
- iv. Rules of probabilities (choose one) **violated** / **obeyed**
 since $0 \leq P(E) \leq 1$, for $E = C, N, D$ and $P(C) + P(N) + P(D) = 1$.
- v. Probability choosing *quarter* is *impossible* since no quarters in box, so
 $P(\text{quarter}) = (\text{choose one}) \mathbf{0} / \frac{1}{10} /$
- vi. Probability choosing cent, nickel or dime is *certain*, so
 $P(C \text{ or } N \text{ or } D) = (\text{choose one}) \mathbf{0} / \frac{1}{10} / \frac{2}{10} / \mathbf{1}$.

5.2 The Addition Rule and Complements

We discuss “and”, “or” and “not” (complement). We look at addition rule for disjoint events, general addition rule and complement rule for probability.

Exercise 5.2 (The Addition Rule and Complements)

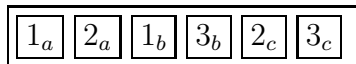
1. *And, Or and Not (Complement)*.

(a) *Summary of operations.*

- *And*: outcomes common to events; in intersection of events
- *Or*: outcomes in union of events
- *Not (Complement)*: outcomes not in event, but in sample space
- *Complement rule*: $P(E^c) = 1 - P(E)$

(b) *Box of tickets.*

Box has six tickets. Each ticket has 1, 2 or 3 with one of three subscripts: a, b or c . One ticket drawn from box at random.



Probability ticket is

- i. “1” is $P(1) = (\text{circle one}) \frac{1}{6} / \frac{2}{6} / \frac{3}{6}$.
- ii. “a” is $P(a) = (\text{circle one}) \frac{1}{6} / \frac{2}{6} / \frac{3}{6}$.
- iii. “1” *and* an “a” is $P(1 \text{ and } a) = (\text{circle one}) \frac{1}{6} / \frac{2}{6} / \frac{3}{6}$.
- iv. “1” *or* an “a” is $P(1 \text{ or } a) = (\text{circle one}) \frac{1}{6} / \frac{2}{6} / \frac{3}{6}$.
- v. “1” *and* a “2” is $P(1 \text{ and } 2) = (\text{circle one}) \frac{0}{6} / \frac{1}{6} / \frac{2}{6}$.
- vi. *not* an “a” is $P(a^c) = 1 - P(a) = 1 - \frac{2}{6} = (\text{circle one}) \frac{2}{6} / \frac{3}{6} / \frac{4}{6}$.

(c) *Box of coins.*

Coins are sampled at random from box.

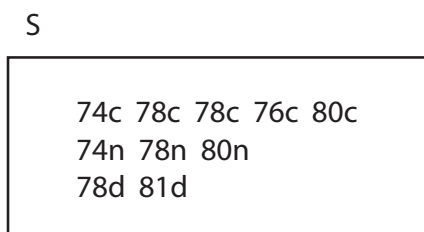


Figure 5.5 (And, or, not: box of coins)

Chance coin is a

- i. cent is $P(C) = (\text{circle one}) \frac{3}{10} / \frac{4}{10} / \frac{5}{10} / \frac{6}{10}$.
- ii. 1978 is $P(1978) = (\text{circle one}) \frac{3}{10} / \frac{4}{10} / \frac{5}{10} / \frac{6}{10}$.
- iii. cent *and* a 1978 is $P(C \text{ and } 1978) = (\text{circle one}) \frac{1}{10} / \frac{2}{10} / \frac{3}{10} / \frac{4}{10}$.
- iv. cent *or* a nickel is $P(C \text{ or } N) = (\text{circle one}) \frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}$.
- v. cent *or* a 1978 is $P(C \text{ or } 1978) = (\text{circle one}) \frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}$.
- vi. *not* a dime is $P(D^c) = (\text{circle one}) \frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}$.

2. *Addition rule.*

(a) Summary.

- Addition rule for *disjoint* events E, F .

$$P(E \text{ or } F) = P(E) + P(F)$$

- *General* addition rule for events E, F .

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

(b) *Box of coins.*

Refer to figure above.

- i. Since not possible to choose a single coin that is both a cent and a nickel; in other words, choosing a cent and nickel are mutually exclusive (disjoint) events, probability of choosing cent *or* a nickel is

$$P(C \text{ or } N) = P(C) + P(N) = \frac{5}{10} + \frac{3}{10} =$$

$$(\text{circle one}) \frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}.$$

- ii. Since it is possible to choose a single coin that is both a cent and a 1978; in other words, choosing a cent and 1978 coin is *not* disjoint, probability of choosing a cent *or* a 1978 is

$$P(C \text{ or } 1978) = P(C) + P(1978) - P(C \text{ and } 1978) = \frac{5}{10} + \frac{4}{10} - \frac{2}{10} =$$

$$\text{(circle one) } \frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}.$$

iii. $P(C \text{ or } D) = P(C) + P(D) = \frac{5}{10} + \frac{2}{10} = \text{(circle one) } \frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}.$

iv. $P(1978 \text{ or } D) = P(1978) + P(D) - P(1978 \text{ and } D) = \frac{4}{10} + \frac{2}{10} - \frac{1}{10} =$
 (circle one) $\frac{5}{10} / \frac{6}{10} / \frac{7}{10} / \frac{8}{10}.$

- (c) *Addition rule: fathers, sons and college.*

Data from a sample of 80 families in a midwestern city gives record of college attendance by fathers (F) and their oldest sons (S).

	son attended college	son did not attend college	
father attended college	18	7	25
father did not attend college	22	33	55
	40	40	80

- i. Probability son, in a randomly chosen family, attended college, is

$$P(S) = \text{(circle one) } \frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{25}{80}.$$

- ii. Probability father, in a randomly chosen family, attended college, is

$$P(F) = \text{(circle one) } \frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{25}{80}.$$

- iii. Probability son *and* father both attended college is

$$P(S \text{ and } F) = \text{(circle one) } \frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{18}{80}.$$

- iv. Probability son *or* father both attended college is

$$P(S \text{ or } F) = P(S) + P(F) - P(S \text{ and } F) = \frac{40}{80} + \frac{25}{80} - \frac{18}{80} =$$

$$\text{(choose one) } \frac{45}{80} / \frac{46}{80} / \frac{47}{80} / \frac{48}{80}.$$

- (d) *More addition rule.*

- i. If $P(E) = \frac{3}{36}$, $P(F) = \frac{9}{36}$ and $P(E \text{ and } F) = \frac{2}{36}$.

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F) = \frac{10}{36} / \frac{11}{36} / \frac{12}{36}.$$

- ii. $P(E) = 0.25$, $P(F) = 0.10$ and $P(E \text{ and } F) = 0.03$,

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F) = \mathbf{0.30} / \mathbf{0.32} / \mathbf{0.33}.$$

- iii. **True / False.** Addition rule determines chance of E “or” F .

- iv. **True / False.** Events E and F are disjoint if $P(E \text{ and } F) = 0$.

5.3 Independence and the Multiplication Rule

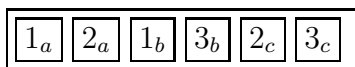
We look at probability of independent events (sampling with replacement) and associated (special case of) multiplication rule: if E , F , G are independent, then

$$P(E \text{ and } F \text{ and } G \text{ and } \dots) = P(E) \cdot P(F) \cdot P(G) \dots$$

Exercise 5.3 (Independence and the Multiplication Rule)1. *Independence (sampling with replacement)*

versus *dependence (sampling without replacement): box of tickets.*

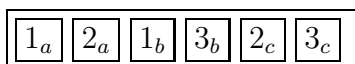
Two things are *independent* if chance for second given first are the same, no matter how first turns out; otherwise, two things are *dependent*.

(a) *Sample with replacement: independence.*

Two tickets are sampled *with* replacement at random from box. All *six* tickets remain in box when second ticket is drawn. Chance second ticket is a “2” *given* first ticket is a “1” is (circle one) $\frac{1}{6} / \frac{2}{6} / \frac{3}{6}$.

(b) When sampling at random with replacement, chance second ticket of two drawn from box is “2”, no matter what the first, is *always* $\frac{1}{6} / \frac{2}{6} / \frac{3}{6}$.(c) *Sample without replacement: dependence.*

Two tickets are sampled *without* replacement at random from box. Only five tickets remain in box when second ticket is drawn. Chance second ticket is a “2” *given* first ticket is a “1” is $\frac{1}{5} / \frac{2}{5} / \frac{3}{5}$.

(d) **True / False** When sampling at random *without* replacement, chance second ticket of two drawn from box is any given number *depends* on number drawn on first ticket.(e) When sampling at random *without* replacement, draws are (circle one) **independent / dependent** of one another; with replacement, draws are independent.(f) **True / False.** If small random samples *without* replacement are taken from large populations (sample size less than 5% of population), it is reasonable to assume independence of events because fractions change so little.2. *Multiplication rule: box of tickets.*(a) Two tickets are sampled with replacement from box: tickets are independent of one another. Chance first ticket is “1” *and* second ticket is “3” is⁴ $P(1 \text{ and } 3) = P(1) \cdot P(3) =$ (circle one) $\frac{1}{6} \times \frac{1}{6} / \frac{2}{6} \times \frac{2}{5} / \frac{2}{6} \times \frac{2}{6} = \frac{1}{9}$. This is an example of (a special case of) the multiplication rule.

⁴Order of tickets matters here. If you did not see how two tickets were picked from box, it would not be clear whether “1” or “3” was first or second ticket and so, in this case, $P(1 \text{ and } 3) + P(3 \text{ and } 1) = P(1) \cdot P(3) + P(3) \cdot P(1) = 2 \times \frac{2}{6} \times \frac{2}{6}$.

- (b) Two tickets are sampled *without* replacement from box: tickets are *dependent* of one another. Chance first ticket is “1” and second ticket is “3” is $P(1 \text{ and } 3) = P(1) \cdot P(3|1) =$ (circle one) $\frac{1}{6} \times \frac{1}{6} / \frac{2}{6} \times \frac{2}{5} / \frac{2}{6} \times \frac{2}{6} = \frac{2}{15}$. This is an example of the (general) multiplication rule.
- (c) **True / False**
 If “1” and “3” independent, then $P(1 \text{ and } 3) = P(1) \cdot P(3)$.
 If “1” and “3” dependent, then $P(1 \text{ and } 3) = P(1) \cdot P(3|1) \neq P(1) \cdot P(3)$.
 In fact, the reverse is true: If $P(1 \text{ and } 3) = P(1) \cdot P(3)$, then “1” and “3” independent, otherwise they are dependent.
- (d) Three tickets are sampled with replacement from box. Chance first ticket is “1” and second ticket is “3” and third ticket is “3” is $P(1 \text{ and } 3 \text{ and } 3) = P(1) \cdot P(3) \cdot P(3) = \frac{2}{6} / \frac{2}{6} \times \frac{2}{6} / \frac{2}{6} \times \frac{2}{6} \times \frac{2}{6} = \frac{1}{27}$.
- (e) Three tickets are sampled with replacement from box. Chance all three tickets are “3”s is⁵ (circle one) $\frac{2}{6} / \frac{2}{6} \times \frac{2}{6} / \frac{2}{6} \times \frac{2}{6} \times \frac{2}{6} = \frac{1}{27}$.
- (f) Three tickets are sampled with replacement from box. Chance *at least one* of the three tickets is a “3” is either
- chance one ticket of three is a “3” or two tickets of three are “3”s or three tickets of three are “3”s, **OR**
 - one minus chance *none* of three tickets are “3”s,
- $$1 - P(\text{no } 3\text{s}) = \text{(circle one)} \quad 1 - \frac{4}{6} / 1 - \frac{4}{6} \times \frac{4}{6} / 1 - \frac{4}{6} \times \frac{4}{6} \times \frac{4}{6} = \frac{19}{27}.$$

3. *Independence versus dependence: fathers, sons and college.*

	son attended college	son did not attend college	
father attended college	18	7	25
father did not attend college	22	33	55
	40	40	80

- (a) Probability son, in a randomly chosen family, attended college, is $P(S) =$ (circle one) $\frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{25}{80}$.
- (b) Probability father, in a randomly chosen family, attended college, is $P(F) =$ (circle one) $\frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{25}{80}$.
- (c) Probability son *and* father both attended college is $P(S \text{ and } F) =$ (circle one) $\frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{18}{80}$.

⁵Order of the tickets does *not* matter here, whether you saw in what order the tickets were chosen or not, since all three tickets are the same: all “3”s. The answer will always be $\frac{2}{6} \times \frac{2}{6} \times \frac{2}{6}$ in this case.

- (d) Since $\frac{18}{80} \neq \left(\frac{40}{80}\right) \times \left(\frac{25}{80}\right)$ or $0.225 \neq 0.15625$; in other words,
 $P(S \text{ and } F) \neq P(S) \times P(F)$,
 event “son attends college”
 (circle one) **is independent of / depends on** event
 “father attends college”.
- (e) **True / False** If $P(A \text{ and } B) = P(A) \cdot P(B)$,
 then events A and B are *independent*, otherwise they are dependent.
 This is one method to determine independence/dependence.

4. *Independence versus disjoint events.* **True / False.**
 Independent events are different from disjoint events.
 Events E and F are independent if $P(E \text{ and } F) = P(E) \cdot P(F)$.
 Events E and F are disjoint if $P(E \text{ and } F) = 0$.

5.4 Conditional Probability and the General Multiplication Rule

We will look at conditional probability

$$P(F|E) = \frac{P(E \text{ and } F)}{P(E)} = \frac{P(F \text{ and } E)}{P(E)},$$

and general multiplication rule,

$$P(E \text{ and } F) = P(E) \cdot P(F|E).$$

Exercise 5.4(Conditional Probability and the General Multiplication Rule)

1. *Conditional probability and dependence: box of coins.*

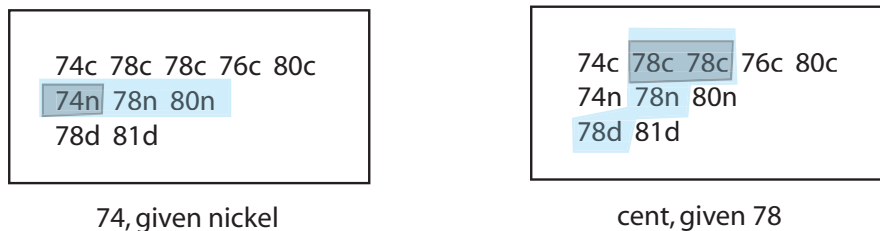


Figure 5.6 (Conditional probability: box of coins)

- (a) *Choosing 1974.*
 Chance a coin chosen at random from box is a 1974 coin is
 $P(1974) =$ (circle one) $\frac{1}{10} / \frac{2}{10} / \frac{3}{10} / \frac{4}{10}$.

(b) *Choosing 1974, given nickel.*

Of three coins that are nickels, (circle one) **1 / 2 / 3** are 1974 coins. Given coin taken from box is a nickel, chance this coin is a 1974 nickel is

$$P(1974|N) = (\text{circle one}) \frac{1}{3} / \frac{2}{3} / \frac{3}{3} / \frac{4}{3}.$$

(c) *Choosing 1974 depends on choosing nickel.*

Unconditional chance coin is "1974", $P(1974) = \frac{2}{10}$, is (circle one) **equal / not equal** to conditional chance coin is "1974, given a nickel", $P(1974|N) = \frac{1}{3}$. Choosing a "1974" and choosing a "nickel" are *dependent*.

(d) *Choosing cent.*

Chance of choosing a cent is $P(C) = (\text{circle one}) \frac{2}{5} / \frac{5}{10} / \frac{2}{10} / \frac{2}{4}$.

(e) *Choosing cent, given 1978.*

Of coins that are 1978s, (circle one) **2 / 4 / 5** are cent coins. Given a coin is a 1978, chance this coin is a cent is $P(C|1978) = \frac{2}{5} / \frac{5}{10} / \frac{2}{10} / \frac{2}{4}$.

(f) *Choosing cent independent of choosing 1978.*

Since $P(C) = \frac{5}{10} = P(C|1978) = \frac{2}{4}$, choosing a "cent" and choosing a "1978" are (choose one) **independent / dependent** events.

(g) *In general.*

If $P(E) = P(E|F)$, events E and F are **dependent / independent**; otherwise, they are *dependent*. This is a second method to determine independence/dependence.

2. More conditional chance: fathers, sons and college.

	son attended college, S	son did not attend college, S^c	
father attended college, F	18	7	25
father did not attend college, F^c	22	33	55
	40	40	80

(a) Probability a son attended college given a father attended college, in a randomly chosen family, is

$$P(S | F) = (\text{circle one}) \frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{25}{80}.$$

(b) Probability a son did not attend college given a father attended college, in a randomly chosen family, is

$$P(S^c | F) = (\text{circle one}) \frac{7}{25} / \frac{18}{25} / \frac{7}{18} / \frac{25}{7}.$$

(c) $P(S | F^c) = (\text{circle one}) \frac{55}{22} / \frac{33}{55} / \frac{22}{55} / \frac{22}{80}$.

(d) $P(S^c | F^c) = (\text{circle one}) \frac{22}{55} / \frac{33}{80} / \frac{22}{33} / \frac{33}{55}$.

(e) $P(F | S) = (\text{circle one}) \frac{18}{40} / \frac{18}{25} / \frac{18}{22} / \frac{25}{80}$
(circle one) **equals / does not equal** $P(S | F) = \frac{18}{25}$.

(f) Using the formula.

$$P(S | F) = \frac{P(S \text{ and } F)}{P(F)} = \frac{18/80}{25/80} =$$

(circle one) $\frac{18}{40} / \frac{18}{25} / \frac{40}{80} / \frac{25}{80}$.

3. *General multiplication rule.* A deck is shuffled and three cards are dealt.

(a) Chance first card dealt is an ace is

$$P(\text{ace}) = (\text{circle one}) \frac{1}{52} / \frac{4}{52} / \frac{3}{51} / \frac{1}{51}.$$

(b) Chance *second* card dealt is a jack, given first card dealt is an ace, is $P(\text{jack} | \text{ace}) = (\text{circle one}) \frac{1}{52} / \frac{4}{50} / \frac{4}{51} / \frac{1}{51}$. This is conditional probability since chance of one event depends on occurrence of another event.

(c) Probability first card is an ace and second card is a jack is

$$P(\text{ace and jack}) = P(\text{ace}) \cdot P(\text{jack} | \text{ace}) =$$

(circle one) $\frac{1}{52} \times \frac{3}{51} / \frac{4}{52} \times \frac{4}{51} / \frac{4}{51} / \frac{1}{51}$.

This is an example of general multiplication rule because it involves product of unconditional probability and conditional probability.

(d) Probability *third* card dealt is a jack, conditional on first two cards dealt are a jack and an ace, is

$$P(\text{jack} | (\text{ace and jack})) = (\text{circle one}) \frac{1}{50} / \frac{4}{52} / \frac{3}{51} / \frac{3}{50}.$$

This is another example of a conditional probability.

(e) Probability of an ace, jack and another jack is

$$P(\text{ace and jack and jack})$$

$$= P(\text{ace}) \cdot P(\text{jack} | \text{ace}) \cdot P(\text{jack} | (\text{ace and jack})) =$$

$$\frac{4}{52} \times \frac{3}{51} \times \frac{2}{50}$$

$$\frac{4}{52} \times \frac{4}{51} \times \frac{3}{50}$$

$$\frac{4}{50} \times \frac{3}{49} \times \frac{2}{48}$$

This is another example of the general multiplication rule⁶.

4. *Addition and Multiplication rules. True / False.*

When “and” is involved, “multiply”: $P(E \text{ and } F) = P(E) \cdot P(F|E)$.

When “or” is involved, “add”: $P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$;

5.5 Counting Techniques

Counting can be an important part of determining various probabilities. To help in understanding various counting techniques, we first discuss the notion of whether

⁶The order of the cards *matters* here. If you did not see how the three cards were picked from the deck, it would not be clear which one of three possibilities occurred: ace, jack, jack or jack, ace, jack or jack, jack, ace. In this case, the answer would be $3 \times \frac{4}{52} \times \frac{4}{51} \times \frac{3}{50}$.

order matters or not. Then, we look at two counting rules: permutations and combinations. Along the way, we find out two ways of visualizing counting techniques: tree diagrams and “marbles in a box”.

Exercise 5.5 (Counting Techniques)

1. *Summary.*

- *Multiplication rule of counting:* If n_1 possible first choices, n_2 second choices and so on, then total of $n_1 \cdot n_2 \cdot \dots$ choices.
- *Permutation:* count of ordered arrangement of r from n distinct objects, sampled without replacement

$${}_n P_r = \frac{n!}{(n-r)!}$$

- *Combination:* count of *unordered* arrangement of r from n distinct objects, sampled without replacement

$${}_n C_r = \frac{n!}{r!(n-r)!}$$

- *Permutation of indistinct items:* count of ordered arrangement of n distinct objects into n_1, n_2, \dots, n_k groups, sampled without replacement

$$\frac{n!}{n_1! n_2! \cdots n_k!}$$

- *Permutation of distinct items with replacement:* count of ordered arrangement of r of n distinct objects, sampled with replacement

$$n^r$$

2. *Order matters or not when counting.*

- Committees (Generic).* Three possible committees of two people from Jim, Sue and Ali are (Jim, Sue), (Jim, Ali) and (Sue, Ali). We (choose one) **would / would not** include both (Jim, Sue) and (Sue, Jim) in the count because *order* of people chosen for this committee *does not matter*.
- Committees (Different Roles).* We (choose one) **would / would not** include (Jim, Sue) and (Sue, Jim) in a count of (chair, secretary) committees because of two roles in committee; *order* of people chosen *does matter*.
- Street numbers.* When counting 3-digit street numbers, (circle one) **order matters / order does not matter**. Are (9,3,4) and (3,4,9) two street numbers (order matters) or one street number (order does not matter)?

- (d) *Cards*. When dealing five cards for a poker hand, (circle one) **order matters / order does not matter**. Are (10,J,Q,2,3) and (2,3,J,Q,10) two hands (order matters) or one hand (order does not matter)?
- (e) *Cars*. When parking cars in parking spots (P1, P2, P3) where P1, P2 and P3 have different parking fees, (circle one) **order matters / order does not matter**. Is parking (Ford,GM,Toyota) same arrangement as parking (GM,Toyota,Ford) in (P1, P2, P3)?
- (f) *Multiple-choice quiz questions*. When answering a sequence of multiple-choice questions on a quiz, (circle one) **order matters / order does not matter**. One possible sequence of answers on a quiz with five questions is A, A, D, A, C and another is D, A, A, C, A, for example.
- (g) We count **more / less** if we assume order matters.

3. *Marbles in a box used for counting outcomes.*

- (a) Number of 3-digit street numbers, as indicated in “marbles in boxes” counting method below: (choose *two!*) $10 \times 10 \times 10$ / 100 / 1000 .

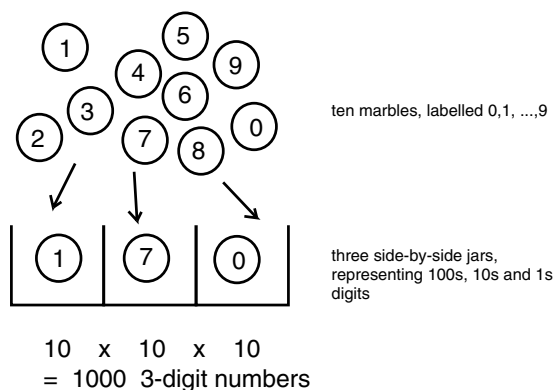


Figure 5.7 (Marbles in boxes: counting 3-Digit street numbers)

- (b) Number of three-digit numbers if first number cannot be zero: (circle *one or more!*) $9 \times 10 \times 10$ / 900 / 999 .
- (c) Number of three-digit numbers if first number must be 3: (circle *one or more!*) $1 \times 10 \times 10$ / 100 / 900 .
- (d) Number three-digit numbers if first number is 3, second cannot be 9: (circle one) $9 \times 9 \times 10$ / $9 \times 10 \times 10$ / $1 \times 9 \times 10$.
- (e) Number of *four*-digit numbers if first number cannot be zero: (circle one) $9 \times 10 \times 10$ / $9 \times 10 \times 10 \times 10$ / $1 \times 9 \times 10 \times 10$.
- (f) Number of three-digit numbers with exactly two 3s: (circle one) $1 \times 1 \times 9 = 9$ / $1 \times 9 \times 1 = 9$ / $9 \times 1 \times 1 = 9$ / $9 + 9 + 9 = 27$.

Hint: Correct answer is fourth one, 27, and is obtained by adding first three possible answers together. Each of first three answers represent a different way to have exactly two 3s in three digits.

4. Tree diagrams used for multiplication rule of counting outcomes.

- (a) As shown in tree diagram, there are (choose one) $3 + 2 = 5$ / $3 \times 2 = 6$ possible (treasurer, secretary) pairs from three eligible treasurer candidates (T_1 , T_2 and T_3) and two eligible secretary candidates (S_1 and S_2), including $\{(T_1, S_1), (T_1, S_2), (T_2, S_1), (T_2, S_2), (T_3, S_1), (T_3, S_2)\}$.

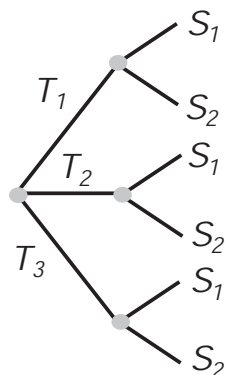


Figure 5.8 (Multiplication rule of counting: treasurer and secretary pairs)

- (b) Number of treasurer and secretary pairs, when *four* eligible treasurer candidates, *eight* eligible secretary candidates, is (circle one) $4 + 8 = 12$ / $4 \times 8 = 24$ / **32**.
- (c) Number of treasurer, secretary and president triplets, when *three* eligible treasurer candidates, *five* eligible secretary candidates and *two* eligible president candidates (circle one) $3 + 5 + 2 = 10$ / **20** / $3 \times 5 \times 2 = 30$.
5. Factorial notation used in formulas for counting outcomes.

- (a) Special mathematical notation, called *factorial notation*, denoted by an exclamation mark, “!”, is often used in formulas used to count outcomes in a sample space. For example,

$$5! = 5 \times 4 \times 3 \times 2 \times 1 =$$

(choose one) **100** / **110** / **120**.

(StatCrunch: Data, Compute expression, fact(5), Compute.)

- (b) 7! is equal to (circle none, one or more)

- i. $7 \times 6!$
- ii. 5040
- iii. $7 \times 6 \times 5!$
- iv. $7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

(StatCrunch: Data, Compute expression, fact(7), Compute.)

- (c) $\frac{7!}{5!}$ is equal to (circle none, one or more)
- $\frac{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{5 \times 4 \times 3 \times 2 \times 1}$
 - 7×6
 - 42
- (d) $\frac{7!}{5!3!}$ is equal to (circle none, one or more)
- $\frac{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{(5 \times 4 \times 3 \times 2 \times 1)(3 \times 2 \times 1)}$
 - $\frac{7 \times 6 \times 5 \times 4}{5 \times 4 \times 3 \times 2 \times 1}$
 - $\frac{7 \times 6}{3 \times 2 \times 1}$
 - $\frac{42}{6}$
- (e) $(7 - 3)!$ is equal to
(circle none, one or more) $7! - 3! / 4! / 4 \times 3 \times 2 \times 1 / 24$.
- (f) $\frac{7!}{(7-3)!}$ is equal to (circle none, one or more)
- $\frac{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{4 \times 3 \times 2 \times 1}$
 - $7 \times 6 \times 5$
 - 210
- (g) *By definition* (in other words, accept as true that), $0! = 1$, and so $0!$ is equal to (circle one) $1! / 2! / 3!$.

6. *Permutation formula (order matters) used in counting outcomes.*

- (a) How many different ways can three of five cars be parked in three different side-by-side parking spots with three different parking fees? As shown in marbles in boxes diagram below, since five different cars occupy first parking spot, only four occupy second parking spot (since one car is in first parking spot) and three could occupy final parking spot, number of permutations is
(circle one) $5 + 4 + 3 = 12 / 5 \times 4 = 20 / 5 \times 4 \times 3 = 60$.

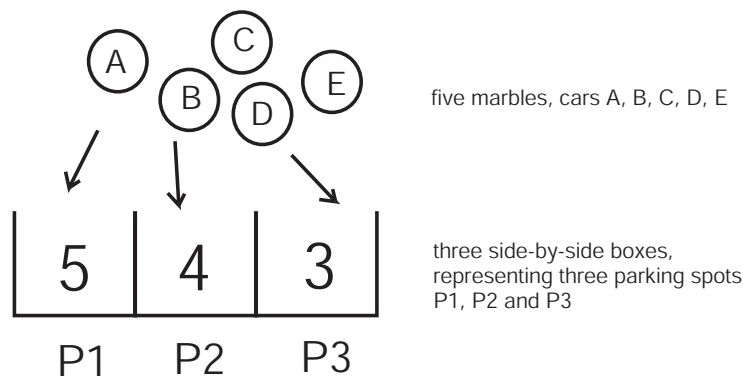


Figure 5.9 (Permutations: arrangements of parked cars)

- (b) **True / False** In marbles in boxes analogy, side-by-side boxes represent parking spots and marbles represent cars, A, B, C, D and E.
- (c) Since (A,B,C) and (B,A,C) count as two different parking arrangements, even though same three cars are used for both, “marbles in boxes” counting method assumes (circle one) **order matters / order does not matter**.
- (d) Number of ways to park 5 cars in 4 parking spots (circle one)
 $5 \times 4 \times 3 \times 2 = 120 / 4 \times 3 \times 2 \times 1 = 24 / 5 \times 4 \times 3 = 120$.
- (e) Number of ways to park $n = 5$ cars in $r = 4$ spots (circle one or more)
- $\frac{5 \times 4 \times 3 \times 2 \times 1}{1}$
 - $\frac{5!}{1!}$
 - ${}_n P_r = \frac{n!}{(n-r)!}$, where $n = 5$ and $r = 4$
- (f) Number of ways to park $n = 7$ cars in $r = 4$ spots (circle one or more)
- ${}_n P_r = \frac{n!}{(n-r)!}$, where $n = 7$ and $r = 4$
 - $\frac{7!}{(7-4)!}$
 - $\frac{7!}{3!}$
 - $7 \times 6 \times 5 \times 4 = 840$
- (StatCrunch: Data, Compute expression, perm(7,4), Compute.)
- (g) Number of ways to park $n = 12$ cars in $r = 8$ spots (circle one or more)
- ${}_n P_r = \frac{n!}{(n-r)!}$, where $n = 12$ and $r = 8$
 - ${}_{12} P_8 = \frac{12!}{(12-8)!}$
 - $\frac{12!}{4!}$
 - $12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 = 19,958,400$

7. *Combination formula (order does not matters) used in counting outcomes.*

- (a) Number of ways of dealing three cards from five cards (10,J,Q,K,A) is calculated by assuming order matters (5 marbles, three side-by-side boxes),

$$5 \times 4 \times 3 = 60$$

and then “dividing out the order” ($3! = 6$ permutations of three cards),

$$\frac{5 \times 4 \times 3}{3!} = \frac{60}{6} =$$

(circle one) **9 / 10 / 11** combinations. See figure below.

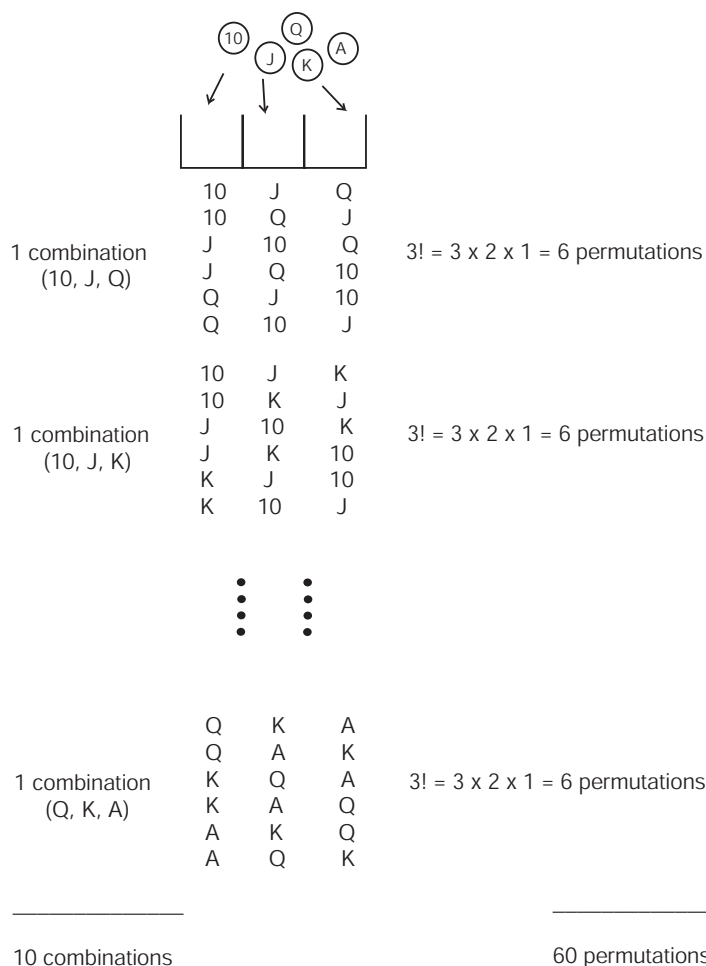


Figure 5.10 (Combinations: number of hands in 5 cards)

(b) Number of ways of dealing $r = 3$ of $n = 5$ cards (choose one or more)

- i. $\frac{5 \times 4 \times 3}{3!}$
- ii. $\frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1)3!}$
- iii. $\frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1)(3 \times 2 \times 1)}$
- iv. ${}_5C_3 = \frac{5!}{2!3!} = 10$
- v. ${}_nC_r = \frac{n!}{(n-r)!r!}$, where $n = 5$ and $r = 3$

(c) Number of ways of dealing $r = 3$ of $n = 11$ cards (choose one or more)

- i. ${}_nC_r = \frac{n!}{(n-r)!r!}$, where $n = 11$ and $r = 3$
- ii. $\frac{11!}{(11-3)!3!}$
- iii. $\frac{11!}{8!3!}$
- iv. $\frac{11 \times 10 \times 9}{3 \times 2 \times 1}$
- v. ${}_{11}C_3 = 165$

(StatCrunch: Data, Compute expression, comb(11,3), Compute.)

8. *Counting and probability: PIN numbers.*

- (a) As shown in figure below, chance four-digit PIN number drawn from a barrel begins with 3 is

$$\frac{1 \times 10 \times 10 \times 10}{10 \times 10 \times 10 \times 10} =$$

(circle one or more) $\frac{1000}{10000}$ / **0.1** / **10%**.

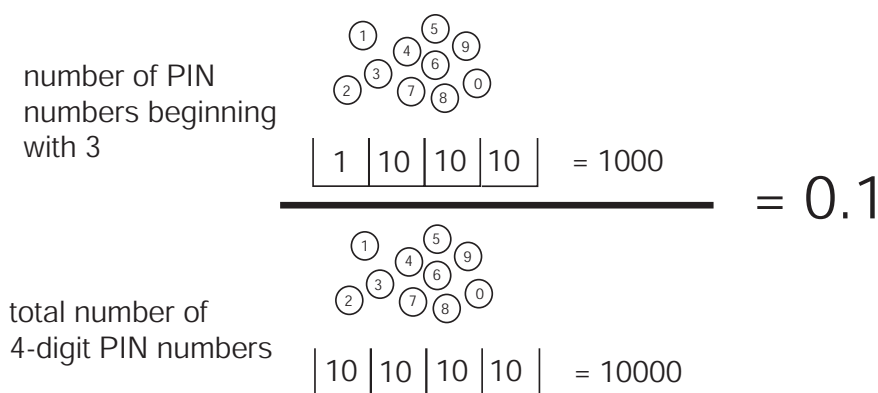


Figure 5.11 (Probability and counting: PIN numbers)

- (b) Chance four-digit PIN number drawn from a barrel has exactly three 3s is $P(3 \text{ threes}) =$ (circle one) $\frac{36}{10000} / \frac{2916}{10000} / \frac{1}{10000} = 0.0036$ or 0.36%

(Look at last part of question 2 above.)

- (c) Chance four-digit PIN number drawn from a barrel has exactly one 3 is $P(\text{one three}) =$ (circle one) $\frac{36}{10000} / \frac{2916}{10000} / \frac{1}{10000} = 0.2916$ or 29.16%

(Again, look at last part of question 2 above.)

- (d) Chance four-digit PIN number drawn from a barrel is the number 1234 is $P(1234) =$ (circle one) $\frac{36}{10000} / \frac{2916}{10000} / \frac{1}{10000} = 0.0001$ or 0.01%

This is *unusual* because $P(1234) = 0.0001$ is small, smaller than 0.05.

9. *Permutations in probability.* Since there are $5 \times 4 \times 3 = 60$ permutations for 3 of five cars, A, B, C, D and E, to park in 3 parking spots, and, furthermore, $1 \times 4 \times 3 = 12$ ways for car A to park in the first spot (and two other cars to park in the other two spots), the chance car A parks in the first spot is $\frac{12}{60} = 0.2$ or 20%. With this in mind, the chance that car A parks in the first spot, if ...

- (a) ... 4 of 5 cars park in 4 parking spots, is

(circle one) $\frac{1 \times 4 \times 3 \times 2}{5 \times 4 \times 3 \times 2} / \frac{5 \times 4 \times 3 \times 2}{5 \times 4 \times 3 \times 2} / \frac{1 \times 3 \times 2 \times 1}{5 \times 4 \times 3 \times 2} = \frac{1 \times 4 P_3}{5 P_4} = 0.20$ or 20%

- (b) ... 4 of 7 cars park in 4 parking spots, is
(circle one) $\frac{1 \times 7 \times 6 \times 5}{5 \times 4 \times 3 \times 2} / \frac{1 \times 6 \times 5 \times 4}{7 \times 6 \times 5 \times 4} / \frac{1 \times 4 \times 3 \times 2}{7 \times 6 \times 5 \times 4} = \frac{1 \times 6 P_3}{7 P_4} \approx 0.1429$ or 14.29%
- (c) ... 4 of 6 cars park in 4 parking spots (remember: car A parks in the first spot) car B must appear in the second spot, is
(circle one) $\frac{1 \times 1 \times 6 \times 5}{6 \times 5 \times 4 \times 3} / \frac{1 \times 1 \times 4 \times 3}{6 \times 5 \times 4 \times 3} / \frac{1 \times 5 \times 4 \times 3}{6 \times 5 \times 4 \times 3} = \frac{1 \times 1 \times 4 P_2}{6 P_4} \approx 0.0333$ or 3.33%
This is *unusual* because 3.33% is small, smaller than 5%.
10. *Combinations in probability.* Since there are ${}_5C_3 = \frac{5!}{3!2!} = 10$ combinations for dealing 3-card hands from five cards, 10, J, Q, K and A, and, furthermore, $1 \times {}_4C_2 = 1 \times \frac{4!}{2!2!} = 6$ ways for a jack to appear in a 3-card hand, chance a jack appears in a 3-card hand is $\frac{1 \times {}_4C_2}{{}_5C_3} = \frac{6}{10} = 0.6$ or 60%.
- (a) If dealt 4 cards from 5 different cards, chance a jack appears:
(circle one) $\frac{1 \times {}_3C_2}{{}_5C_4} / \frac{{}_6C_3}{{}_7C_4} / \frac{1 \times {}_4C_3}{{}_5C_4} / \frac{1 \times {}_4C_3}{{}_4C_2} \approx 0.80$ or 80%
- (b) If dealt 4 cards from 7 different cards, chance a jack appears:
(circle one) $\frac{1 \times {}_4C_{4,3}}{{}_5C_4} / \frac{1 \times {}_6C_4}{{}_7C_4} / \frac{1 \times {}_6C_3}{{}_7C_4} / \frac{1 \times {}_3C_6}{{}_7C_4} \approx 0.5714$ or 57.14%
- (c) If dealt 4 cards from 6 different cards, chance a jack and ten appears:
(circle two) $\frac{1 \times 1 \times {}_4C_2}{{}_6C_4} / \frac{{}_6C_3}{{}_6C_4} / \frac{1 \times 1 \times {}_5C_3}{{}_6C_4} / \frac{{}_4C_3}{{}_5C_4} / \frac{{}_4C_2}{{}_6C_4} \approx 0.40$ or 40%
11. *Another combination and probability question: choosing people.* Determine chance Scout, Buzz and Mary and two others are chosen from 12 people.
- (a) Order (circle one) **matters / does not matter.**
- (b) Total number of ways 5 people can be chosen from 12 is
(circle three) $\frac{12!}{5!} / {}_{12}C_7 / \frac{12!}{5!7!} / {}_{12}C_5$
- (c) Chance Scout, Buzz and Mary and two others are chosen from 12 is
(circle one) $\frac{12!9!}{5!} / \frac{1 \times 1 \times 1 \times 2!}{{}_{12}C_7} / \frac{12!9!}{5!7!} / \frac{1 \times 1 \times 1 \times {}_9C_2}{{}_{12}C_5} \approx 0.0455$ or 4.55%
This is *unusual* because 4.55% is small, smaller than 5%.
12. *Another probability question using counting: multiple choice quiz.* On first question on a multiple choice quiz, of four possible choices (A, B, C and D), answer C might be correct, say. One possible answer key to the six questions on this quiz might be C, B, A, D, A and C, say. If an answer key is created at random, determine chance first two answers to this answer key are both A.
- (a) Order (circle one) **matters / does not matter.**
- (b) “Marbles” are (circle one) **six questions / four answers** and “boxes” are (circle one) **six questions / four answers.**
- (c) If marbles are four answers and boxes are six questions, total number of possible different answer keys is (circle one) $4^6 / 6^4 / {}_6P_4 / {}_6C_4 = 4096$
- (d) Chance first two answers to this answer key are both A is
(circle one) $\frac{4^6}{6^4} / \frac{1 \times 1 \times 6^4}{4^6} / \frac{{}_6P_4}{{}_6C_4} / \frac{1 \times 1 \times 6 C_4}{4^6} / \frac{1 \times 1 \times 4^4}{4^6} \approx 0.0625$ or 6.25%

5.6 Putting It Together: Which Method Do I Use?

We will skip this section because we have covered everything in it. This section summarizes all material in previous sections.