UCLA STAT 110 A
Applied Probability & Statistics for Engineers

- **Instructor:** Ivo Dinov,
  Asst. Prof. In Statistics and Neurology
- **Teaching Assistant:** Maria Chang, UCLA Statistics

University of California, Los Angeles, Spring 2003

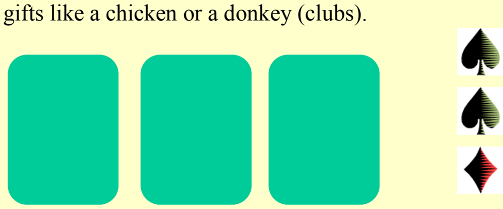
http://www.stat.ucla.edu/~dinov/

Chapter 2: Probability

- Where do probabilities come from?
- Simple probability models
- probability rules
- Conditional probability
- Statistical independence

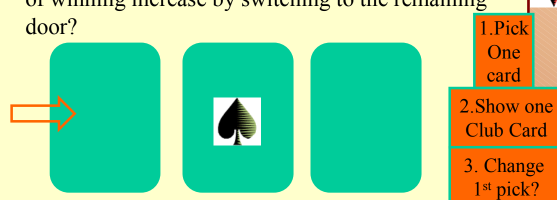
Let's Make a Deal Paradox – aka, Monty Hall 3-door problem

- This paradox is related to a popular television show in the 1970's. In the show, a contestant was given a choice of three doors/cards of which one contained a prize (diamond). The other two doors contained gag gifts like a chicken or a donkey (clubs).



Let's Make a Deal Paradox.

- After the contestant chose an initial door, the host of the show then revealed an empty door among the two unchosen doors, and asks the contestant if he or she would like to switch to the other unchosen door. The question is should the contestant switch. Do the odds of winning increase by switching to the remaining door?



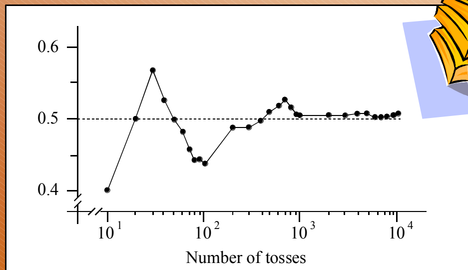
Let's Make a Deal Paradox.

- The *intuition* of most people tells them that each of the doors, the chosen door and the unchosen door, are equally likely to contain the prize so that there is a 50-50 chance of winning with either selection? This, however, is **not the case**.
- The probability of winning by using the switching technique is 2/3, while the odds of winning by not switching is 1/3. The easiest way to explain this is as follows:

Let's Make a Deal Paradox.

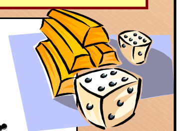
- The probability of picking the wrong door in the initial stage of the game is 2/3.
- If the contestant picks the wrong door initially, the host must reveal the remaining empty door in the second stage of the game. Thus, if the contestant switches after picking the wrong door initially, the contestant will win the prize.
- The probability of winning by switching then reduces to the probability of picking the wrong door in the initial stage which is clearly 2/3.

Long run behavior of coin tossing



Proportion of heads versus number of tosses for John Kerrich's coin tossing experiment.

From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.



Definitions …

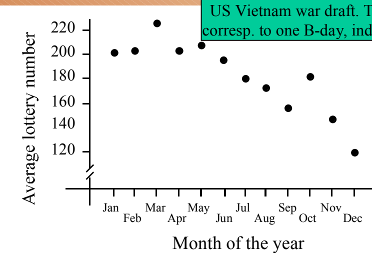
- The law of averages about the behavior of coin tosses – the relative proportion (relative frequency) of heads-to-tails in a coin toss experiment becomes more and more stable as the number of tosses increases. The law of averages applies to relative frequencies not absolute counts of #H and #T.
- Two widely held misconceptions about what the law of averages about coin tosses:
  - Differences between the actual numbers of heads & tails becomes more and more variable with increase of the number of tosses – a seq. of 10 heads doesn't increase the chance of a tail on the next trial.
  - Coin toss results are fair, but behavior is still unpredictable.

Coin Toss Models

- Is the coin tossing model adequate for describing the sex order of children in families?
  - This is a rough model which is not exact. In most countries rates of B/G is different; form 48% … 52%, usually. Birth rates of boys in some places are higher than girls, however, female population seems to be about 51%.
  - Independence, if a second child is born the chance it has the same gender (as the first child) is slightly bigger.

Data from a "random" draw

366 cylinders (for each day in the year) for the US Vietnam war draft. The N-th drawn number, corresp. to one B-day, indicating order of drafting.



So, people born later in the year tend to have lower lottery numbers and a bigger chance of actually being drafted.

Average lottery numbers by month.
Replotted from data in Fienberg [1971].

From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.

Types of Probability

- Probability models have two essential components (*sample space*, the space of all possible outcomes from an experiment; and a list of *probabilities* for each event in the sample space). Where do the outcomes and the probabilities come from?
- Probabilities from models – say mathematical/physical description of the sample space and the chance of each event. Construct a fair die tossing game.
- Probabilities from data – data observations determine our probability distribution. Say we toss a coin 100 times and the observed Head/Tail counts are used as probabilities.
- Subjective Probabilities – combining data and psychological factors to design a reasonable probability table (e.g., gambling, stock market).

Sample Spaces and Probabilities

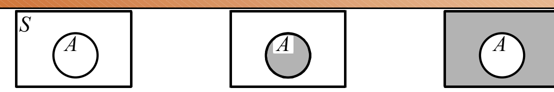
- When the relative frequency of an event in the past is used to estimate the probability that it will occur in the future, what assumption is being made?
  - The underlying process is stable over time;
  - Our relative frequencies must be taken from large numbers for us to have confidence in them as probabilities.
- All statisticians agree about how probabilities are to be combined and manipulated (in math terms), however, not all agree what probabilities should be associated with a particular real-world event.
- When a weather forecaster says that there is a 70% chance of rain tomorrow, what do you think this statement means? (Based on our past knowledge, according to the barometric pressure, temperature, etc. of the conditions we expect tomorrow, 70% of the time it did rain under such conditions.)

Sample spaces and events

- A ***sample space***, *S*, for a random experiment is the set of all possible outcomes of the experiment.
- An ***event*** is a *collection* of outcomes.
- An event ***occurs*** if any outcome making up that event occurs.

The complement of an event

- The complement of an event *A*, denoted $\bar{A}$, occurs *if and only if A* does not occur.

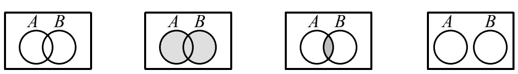


(a) Sample space containing event *A* (b) Event *A* shaded (c) $\bar{A}$ shaded

An event *A* in the sample space *S*.

Combining events – *all statisticians agree on*

- ***"A* or *B"*** contains all outcomes in *A* **or** *B* (or both).
- ***"A* and *B"*** contains all outcomes which are **in both** *A* and *B*.



(a) Events *A* and *B* (b) "*A* or *B*" shaded (c) "*A* and *B*" shaded (d) Mutually exclusive events

Two events.

Mutually exclusive events cannot occur at the same time.

Probability distributions

- Probabilities always lie between 0 and 1 and they sum up to 1 (across all simple events) .
- ***pr(A)*** can be obtained by adding up the probabilities of all the outcomes in *A*.

$$pr(A) = \sum_{\substack{E \text{ outcome} \\ \text{in event } A}} pr(E)$$

Review

- Law of averages for the coin-toss example.
- Sample spaces, outcomes, events, complements.
- Probabilities are always in the range [0 : 1]
- ***pr(A)*** can be obtained by adding up the probabilities of all the outcomes in *A*.

$$pr(A) = \sum_{\substack{E \text{ outcome} \\ \text{in event } A}} pr(E)$$

Job losses in the US

Job Losses in the US (in thousands)
for 1987 to 1991

	Reason for Job Loss			
	Workplace moved/closed	**Slack work**	**Position abolished**	**Total**
Male	1,703	1,196	548	3,447
Female	1,210	564	363	2,137
Total	2,913	1,760	911	5,584

Job losses cont.

	Workplace moved/closed	Slack work	Position abolished	Total
Male	1,703	1,196	548	3,447
Female	1,210	564	363	2,137
Total	2,913	1,760	911	5,584

Proportions of Job Losses from Table 4.4.1

	Reason for Job Loss			
	Workplace moved/closed	Slack work	Position abolished	Row totals
Male	.305	.214	.098	.617
Female	.217	.101	.065	.383
Column totals	.552	.315	.163	1.000

Review

- What is a sample space? What are the <u>two essential criteria</u> that must be satisfied by a possible sample space? (completeness – every outcome is represented; and uniqueness – no outcome is represented more than once.
- What is an event? (collection of outcomes)
- If A is an event, what do we mean by its complement, $\bar{A}$? When does $\bar{A}$ occur?
- If A and B are events, when does ***A* or *B*** occur? When does ***A* and *B*** occur?

Properties of probability distributions

- A sequence of number $\{p_1, p_2, p_3, \ldots, p_n\}$ is a probability distribution for a sample space $S = \{s_1, s_2, s_3, \ldots, s_n\}$, if $pr(s_k) = p_k$, for each $1<=k<=n$. The two essential properties of a probability distribution $p_1, p_2, \ldots, p_n$?

$$p_k \geq 0; \quad \sum_k p_k = 1$$

- How do we get the probability of an event from the probabilities of outcomes that make up that event?
- If all outcomes are <u>distinct</u> & <u>equally likely</u>, how do we calculate *pr(A)* ? If $A = \{a_1, a_2, a_3, \ldots, a_9\}$ and $pr(a_1)=pr(a_2)=\ldots=pr(a_9)=p$; then

$$\underline{pr(A) = 9 \times pr(a_1) = 9p}.$$

Example of probability distributions

- Tossing a coin twice. *Sample space* S={HH, HT, TH, TT}, for a fair coin each outcome is equally likely, so the probabilities of the 4 possible outcomes should be identical, *p*. Since, p(HH)=p(HT)=p(TH)=p(TT)=p and

$$p_k \geq 0; \quad \sum_k p_k = 1$$

- p = ¼ = 0.25.

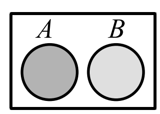
Proportion vs. Probability

- How do the concepts of a proportion and a probability <u>differ</u>? A proportion is a <u>partial description</u> of a real population. The probabilities give us the <u>chance</u> of something happening in a random experiment. Sometimes, proportions are <u>identical</u> to probabilities (e.g., in a real population under the experiment *choose-a-unit-at-random)*.
- See the ***two-way table of counts*** (*contingency table*) on Table 4.4.1, slide 19. E.g., *choose-a-person-at-random* from the ones laid off, and compute the chance that the person would be a <u>male</u>, laid off due to <u>position-closing</u>. We can apply the same rules for manipulating probabilities to proportions, in the case where these two are identical.

Rules for manipulating Probability Distributions

For mutually exclusive events,

pr(*A* or *B*) = pr(*A*) + pr(*B*)



From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.

Unmarried couples

Select an unmarried couple *at random* – the table proportions give us the probabilities of the events defined in the row/column titles.

Proportions of Unmarried Male-Female Couples Sharing Household in the US, 1991

Male	Female: Never Married	Divorced	Widowed	Married to other	Total
Never Married	0.401	.111	.017	.025	.554
Divorced	.117	.195	.024	.017	.353
Widowed	.006	.008	.016	.001	.031
Married to other	.021	.022	.003	.016	.062
Total	.545	.336	.060	.059	1.000

Review

- If A and B are mutually exclusive, what is the probability that both occur? (0) What is the probability that at least one occurs? (sum of probabilities)
- If we have two or more mutually exclusive events, how do we find the probability that at least one of them occurs? (sum of probabilities)
- Why is it sometimes easier to compute $pr(A)$ from $pr(A) = 1 - pr(\bar{A})$? (The complement of the even may be easer to find or may have a known probability. E.g., a random number between 1 and 10 is drawn. Let A ={a number less than or equal to 9 appears}. Find $pr(A) = 1 - pr(\bar{A})$). probability of $\bar{A}$ is pr({10 appears}) = 1/10 = 0.1. Also Monty Hall 3 door example!

Melanoma – type of skin cancer – an example of laws of conditional probabilities

400 Melanoma Patients by Type and Site

Type	Site: Head and Neck	Trunk	Extremities	Row Totals
Hutchinson's melanomic freckle	22	2	10	34
Superficial	16	54	115	185
Nodular	19	33	73	125
Indeterminant	11	17	28	56
Column Totals	68	106	226	400

Contingency table based on Melanoma histological type and its location

Conditional Probability

The ***conditional probability*** of ***A*** occurring ***given*** that *B* occurs is given by

$$\text{pr}(A \mid B) = \frac{\text{pr}(A \text{ and } B)}{\text{pr}(B)}$$

Suppose we select one out of the 400 patients in the study and we want to find the probability that the cancer is on the *extremities* *given that* it is of type *nodular*: P = 73/125 = P(C. on Extremities | Nodular)

$$\frac{\#\text{nodular patients with cancer on extremities}}{\#\text{nodular patients}}$$

Multiplication rule- what's the percentage of Israelis that are poor and Arabic?

$$\text{pr}(A \text{ and } B) = \text{pr}(A \mid B)\text{pr}(B) = \text{pr}(B \mid A)\text{pr}(A)$$

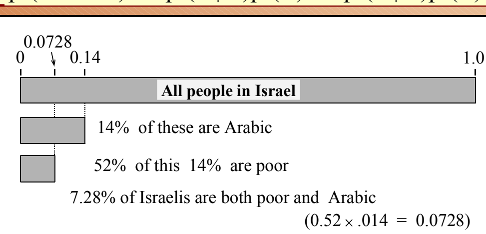


Illustration of the multiplication rule.

From *Chance Encounters* by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000.

Review

$$\text{pr}(A \text{ and } B) = \text{pr}(A \mid B)\text{pr}(B) = \text{pr}(B \mid A)\text{pr}(A)$$

$$\text{pr}(A) = 1 - \text{pr}(\bar{A})$$

1. Proportions (partial description of a real population) and probabilities (giving the chance of something happening in a random experiment) may be identical – under the experiment *choose-a-unit-at-random*
2. Properties of probabilities.

$$\{p_k\}_{k=1}^{N} \text{ define probabilities} \Leftrightarrow p_k \geq 0; \ \sum_k p_k = 1$$

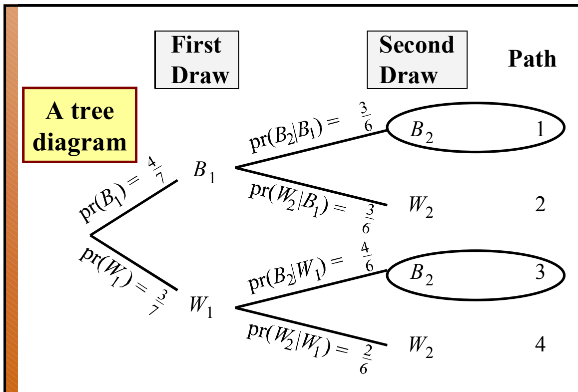
A tree diagram for computing conditional probabilities

Suppose we draw 2 balls at random one at a time *without replacement* from an urn containing **4 black** and **3 white** balls, otherwise identical. What is the probability that the *second ball is black*? Sample Spc?

P({*2-nd ball is black*}) =
P({*2-nd is black*} &{*1-st is black*}) +
P({2-nd is black} &{*1-st is white*}) =
4/7 x 3/6 + 4/6 x 3/7 = 4/7.

Mutually exclusive

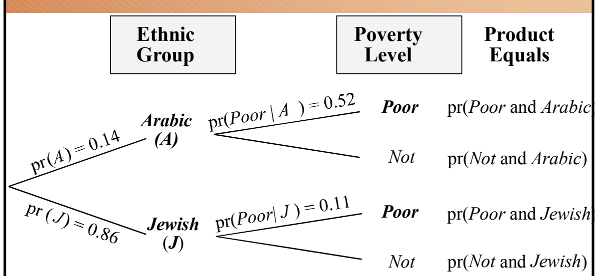
Slide 31 STAT 110A, UCLA, Ivo Dinov



Tree diagram for a sampling problem.

Slide 32 STAT 110A, UCLA, Ivo Dinov

Tree diagram for poverty in Israel



Slide 33 STAT 110A, UCLA, Ivo Dinov

2-way table for poverty in Israel

pr(*Poor* and *Arabic*) = pr(*Poor*|*Arabic*) × pr(*Arabic*) [= 52% of 14%]

pr(*Poor* and *Jewish*) = pr(*Poor*|*Jewish*) × pr(*Jewish*) [= 11% of 86%]

		Ethnicity: Arabic	Ethnicity: Jewish	Total
Poverty	Poor	.52 × .14	.11 × .86	?
	Not poor	?	?	?
	Total	.14	.86	1.00

pr(*Arabic*) = .14 pr(*Jewish*) = .86

Proportions by Ethnicity and Poverty.

P(A & B) = P(A | B) x P(B),
P(A | B) = P(A & B) / P(B)
P(A & B) = P(B & A) = P(B | A) x P(A).
P(A | B) = [P(B | A) x P(A)] / P(B).

Slide 34 STAT 110A, UCLA, Ivo Dinov

2-way table for poverty in Israel cont.

pr(*Poor* and *Arabic*) = pr(*Poor*|*Arabic*) × pr(*Arabic*) [= 52% of 14%]

pr(*Poor* and *Jewish*) = pr(*Poor*|*Jewish*) × pr(*Jewish*) [= 11% of 86%]

		Ethnicity: Arabic	Ethnicity: Jewish	Total
Poverty	Poor	.52 × .14	.11 × .86	?
	Not poor	?	?	?
	Total	.14	.86	1.00

pr(*Arabic*) = .14 pr(*Jewish*) = .86

Proportions by Ethnicity and Poverty

		Ethnicity: Arabic	Ethnicity: Jewish	Total
Poverty	Poor	.0728	.0946	.1674
	Not Poor	.0672	.7654	.8326
	Total	.14	.86	1.00

Slide 35 STAT 110A, UCLA, Ivo Dinov

Conditional probabilities and 2-way tables

- Many problems involving conditional probabilities can be solved by constructing two-way tables
- This includes *reversing the order of conditioning*

P(A & B) = P(A | B) x P(B) = P(B | A) x P(A)

Slide 36 STAT 110A, UCLA, Ivo Dinov

Proportional usage of *oral contraceptives* and their rates of failure

We need to complete the two-way contingency table of proportions

pr(*Failed* and *Oral*) = pr(*Failed* | *Oral*) × pr(*Oral*) [= 5% of 32%]

pr(*Failed* and *IUD*) = pr(*Failed* | *IUD*) × pr(*IUD*) [= 6% of 3%]

		Method					
		Steril.	Oral	Barrier	IUD	Sperm.	Total
Outcome	Failed	0 × .38	.05 × .32	.14 × .24	.06 × .03	.26 × .03	?
	Didn't	?	?	?	?	?	?
	Total	.38	.32	.24	.03	.03	1.00

pr(*Steril.*) = .38 pr(*Barrier*) = .24 pr(*IUD*) = .03

Oral contraceptives cont.

pr(*Failed* and *Oral*) = pr(*Failed* | *Oral*) × pr(*Oral*) [= 5% of 32%]

pr(*Failed* and *IUD*) = pr(*Failed* | *IUD*) × pr(*IUD*) [= 6% of 3%]

		Method					
		Steril.	Oral	Barrier	IUD	Sperm.	Total
Outcome	Failed	0 × .38	.05 × .32	.14 × .24	.06 × .03	.26 × .03	?
	Didn't	?	?	?	?	?	?
	Total	.38	.32	.24	.03	.03	1.00

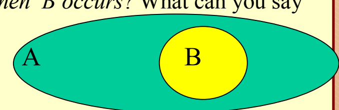
pr(*Steril.*) = .38 pr(*Barrier*) = .24 pr(*IUD*) = .03

Table Constructed from the Data in Example 4.6.

		Method					
		Steril.	**Oral**	**Barrier**	**IUD**	**Sperm.**	**Total**
Outcome	Failed	0	.0160	.0336	.0018	.0078	.0592
	Didn't	.3800	.3040	.2064	.0282	.0222	.9408
	Total	.3800	.3200	.2400	.0300	.0300	1.0000

Remarks …

- In *pr(A | B)*, how should the symbol " | " is read *given that*.
- How do we interpret the fact that: *The event A always occurs when B occurs*? What can you say about pr(*A* | *B*)?



- When drawing a probability tree for a particular problem, how do you know *what events* to use for the first fan of branches and which events to use for the subsequent branching? (at each branching stage condition on all the info available up to here. E.g., at first branching use all simple events, no prior is available. At 3-rd branching condition of the previous 2 events, etc.).

Number of Individuals Having a Given Mean Absorbance Ratio (MAR) in the ELISA for HIV Antibodies

MAR	Healthy Donor	HIV patients
<2	202 } 275	0 } 2 False-Negatives (FNE)
2 - 2.99	73	2
Test cut-off		
3 - 3.99	15 } False-positives	7
4 - 4.99	3	7
5 - 5.99	2	15
6 - 11.99	2	36
12+	0	21
Total	297	88

Power of a test is: 1-P(FNE)= 1-P(Neg|HIV) ~ 0.976

Adapted from Weiss et al.[1985]

HIV cont.

pr(*HIV* and *Positive*) = pr(*Positive*|*HIV*) × pr(*HIV*) [= 98% of 1%]

pr(*Not HIV* and *Negative*) = pr(*Negative*|*Not HIV*) × pr(*Not HIV*) [= 93% of 99%]

		Test result		
		Positive	Negative	Total
Disease status	HIV	.98 × .01	?	.01 ⟵ pr(*HIV*) = .01
	Not HIV	?	.93 × .99	.99 ⟵ pr(*Not HIV*) = .99
	Total	?	?	1.00

Putting HIV information into the table.

HIV – reconstructing the contingency table

pr(*HIV* and *Positive*) = pr(*Positive*|*HIV*) × pr(*HIV*) [= 98% of 1%]

pr(*Not HIV* and *Negative*) = pr(*Negative*|*Not HIV*) × pr(*Not HIV*) [= 93% of 99%]

		Test result		
		Positive	Negative	Total
Disease status	HIV	.98 × .01	?	.01 ⟵ pr(*HIV*) = .01
	Not HIV	?	.93 × .99	.99 ⟵ pr(*Not HIV*) = .99
	Total	?	?	1.00

Proportions by Disease Status and Test Result

		Test Result		
		Positive	**Negative**	**Total**
Disease	HIV	.0098	.0002	.01
Status	Not HIV	.0693	.9207	.99
	Total	.0791	.9209	1.00

Proportions of HIV infections by country

Proportions Infected with HIV

Country	No. AIDS Cases	Population (millions)	pr(HIV)	pr(HIV | Positive)
United States	218,301	252.7	0.00864	0.109
Canada	6,116	26.7	0.00229	0.031
Australia	3,238	16.8	0.00193	0.026
New Zealand	323	3.4	0.00095	0.013
United Kingdom	5,451	57.3	0.00095	0.013
Ireland	142	3.6	0.00039	0.005

Statistical independence

- Events A and B are *statistically independent* if knowing whether B has occurred gives no new information about the chances of A occurring,

  i.e. if **pr(*A* | *B*) = pr(*A*)**

- Similarly, P(B | A) = P(B), since

  P(B|A)=P(B & A)/P(A) = P(A|B)P(B)/P(A) = P(B)

- If A and B are *statistically independent*, then

$$\text{pr}(A \text{ and } B) = \text{pr}(A) \times \text{pr}(B)$$

Example using independence

There are many genetically based blood group systems. Two of these are:
Rh blood type system (Rh+ and Rh-) and the *Kell* system (K+ and K-).
For Europeans the following proportions are experimentally obtained.

Blood Type Data

$.08 \times .81$ · $.92 \times .81$ · pr($RH+$) = .81 · $.08 \times .19$ · $.92 \times .19$ · pr($K+$) = .08

	$K+$	$K-$	Total
$Rh+$	?	?	.81
$Rh-$	?	?	.19
Total	.08	.92	1.00

→

	$K+$	$K-$	Total
$Rh+$	.0648	.7452	.81
$Rh-$	.0152	.1748	.19
Total	.08	.92	1.00

How can we fill in the inside of the two-way contingency table?
It is known that anyone's blood type in one system is *independent* of their type in another system.

P(Rh+ and K+) = P(Rh+) x P(K+)= 0.81 x 0.08 = 0.0648

People vs. Collins

Frequencies Assumed by the Prosecution

Yellow car	$\frac{1}{10}$	Girl with blond hair	$\frac{1}{3}$
Man with mustache	$\frac{1}{4}$	Black man with beard	$\frac{1}{10}$
Girl with ponytail	$\frac{1}{10}$	Interracial couple in car	$\frac{1}{1000}$

- The first occasion where a conviction was made in an American court of law, largely on statistical evidence, 1964. A woman was mugged and the offender was described as a wearing dark cloths, with blond hair in a pony tail who got into a yellow car driven by a black male accomplice with mustache and beard. The suspect brought to trial were picked out in a line-up and fit all of the descriptions. Using the *product rule for probabilities* an expert witness computed the chance that a random couple meets these characteristics, as 1:12,000,000.

Summary

- What does it mean for two events A and B to be *statistically independent*?
- Why is the working rule under independence, *P(A and B) = P(A) P(B)* , just a special case of the multiplication rule *P(A & B) = P(A | B) P(B)* ?
- *Mutual independence* of events $A_1, A_2, A_3, \ldots, A_n$ *if and only* if $P(A_1 \& A_2 \& \ldots \& A_n) = P(A_1)P(A_2)\ldots P(A_n)$
- What do we mean when we say two human characteristics are *positively associated*? *negatively associated*? (blond hair – blue eyes, pos.; black hair – blue eyes, neg.assoc.)

Review

- What happens to the calculated P(A and B) if we treat positively associated events as independent? if we treat negatively associated events as independent?

  (Example, let B={A + {b}}, A & B are pos-assoc'd, P(A&B)=P(A)[P(A)+P({b})], under indep. assump's. However, P(A&B)=P(B|A)P(A)=1 x P(A) **>** P(A)[P(A)+P({b})], underestimating the real chance of events. If A & B are neg-assoc'd → A & comp(B) are pos-assoc'd. In general, this may lead to answers that are grossly too small or too large …)

- Why do people often treat events as independent? When can we trust their answers?(Easy computations! Not always!)

Summary of ideas

- The ***probabilities*** people quote come from 3 main sources:
  - (i) ***Models*** (idealizations such as the notion of equally likely outcomes which suggest probabilities by symmetry).
  - (ii) ***Data*** (e.g.relative frequencies with which the event has occurred in the past).
  - (iii) ***subjective feelings*** representing a degree of belief
- A simple probability model consists of a sample space and a probability distribution.
- A sample space, *S*, for a random experiment is the set of all possible outcomes of the experiment.

Summary of ideas cont.

- A list of numbers $p_1, p_2, \ldots$ is a ***probability distribution*** for $S = \{s_1, s_2, s_3, \ldots\}$, provided
  - all of the p_i's lie between 0 and 1, and
  - they add to 1.
- According to the probability model, p_i is the probability that outcome s_i occurs.
- We write $p_i = P(s_i)$.

Summary of ideas cont.

- An ***event*** is a collection of outcomes
- An event ***occurs*** if any outcome making up that event occurs
- The probability of event A can be obtained by adding up the probabilities of all the outcomes in A
- If all outcomes are equally likely,

$$\text{pr}(A) = \frac{\text{number of outcomes in } A}{\text{total number of outcomes}}$$

Summary of ideas cont.

- The ***complement*** of an event *A*, denoted $\overline{A}$, occurs if *A* does not occur
- It is useful to represent events diagrammatically using ***Venn diagrams***
- ***A union*** of events, ***A* or *B*** contains all outcomes in *A* or *B* (including those in both). It occurs if at least one of *A* or *B* occurs
- ***An intersection*** of events, ***A and B*** contains all outcomes which are in ***both*** *A* and *B*. It occurs only if both *A* and *B* occur
- ***Mutually exclusive*** events cannot occur at the same time

Summary of ideas cont.

- The ***conditional probability*** of *A* occurring ***given*** that *B* occurs is given by

$$\text{pr}(A \mid B) = \frac{\text{pr}(A \text{ and } B)}{\text{pr}(B)}$$

- Events *A* and *B* are ***statistically independent*** if knowing whether *B* has occurred gives no new information about the chances of *A* occurring, i.e. if $P(A \mid B) = P(A)$ → $P(B|A)=P(B)$.
- If events are physically independent, then, under any sensible probability model, they are also statistically independent
- Assuming that events are independent when in reality they are not can often lead to answers that are grossly too big or grossly too small

Formula Summary

- For discrete sample spaces, pr(*A*) can be obtained by adding the probabilities of all outcomes in *A*
- For equally likely outcomes in a finite sample space

$$\text{pr}(A) = \frac{\text{number of outcomes in } A}{\text{total number of outcomes}}$$

Formula summary cont.

- pr(S) = 1
- pr($\bar{A}$) = 1 - pr(A)
- If A and B are mutually exclusive events, then pr(A or B) = pr(A) + pr(B)

  (here "or" is used in the inclusive sense)
- If $A_1, A_2, \ldots, A_k$ are mutually exclusive events, then

  pr(A_1 or A_2 or … or A_k) = pr(A_1)+pr(A_2)+…+pr(A_k)

Formula summary cont.

Conditional probability

- Definition:

$$\text{pr}(A \mid B) = \frac{\text{pr}(A \text{ and } B)}{\text{pr}(B)}$$

- Multiplication formula:

$$\text{pr}(A \text{ and } B) = \text{pr}(B|A)\text{pr}(A) = \text{pr}(A|B)\text{pr}(B)$$

Formula summary cont.

Multiplication Rule under independence:

- If A and B are independent events, then

$$\text{pr}(A \text{ and } B) = \text{pr}(A)\,\text{pr}(B)$$

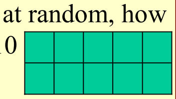
- If $A_1, A_2, \ldots, A_n$ are mutually independent,

pr(A_1 and A_2 and ... and A_n) = pr(A_1) pr(A_2) … pr(A_n)

Theory of Counting = Combinatorial Analysis

Principle of Counting: If 2 experiments are performed and the first one has N_1 possible outcomes, the second (independent) experiment has N_2 possible outcomes then the number of outcomes of the combined (dual) experiment is $N_1 \times N_2$.

E.g., Suppose we have 5 math majors in the class, each carrying 2 textbooks with them. If I select a math major student and 1 textbook at random, how many possibilities are there? 5x2=10



Theory of Counting = Combinatorial Analysis

Generalized Principle of Counting: If M (independent) experiments are performed and the first one has N_m possible outcomes, 1<=m<=M, then the TOTAL number of outcomes of the combined experiment is

$$N_1 x N_2 x \ldots x N_M.$$

E.g., How many binary functions [f(i)=0 or f(i)=1], defined on a grid 1, 2, 3, …, n, are there? How many numbers can be stored in 8 bits = 1 byte?

$$2 \times 2 \times \ldots \times 2 = 2^n$$

Permutation & Combination

Permutation: Number of **ordered** arrangements of **r** objects chosen from **n** ***distinctive*** objects

$$P_n^r = n(n-1)(n-2)\ldots(n-r+1)$$

$$P_n^n = P_n^{n-r} \cdot P_r^r$$

e.g. $P_6^3 = 6 \cdot 5 \cdot 4 = 120$.

Permutation & Combination

Combination: Number of **non-ordered** arrangements of r objects chosen from n *distinctive* objects: $C_n^r = P_n^r / r! = \frac{n!}{(n-r)!r!}$

Or use notation of $\binom{n}{r} = C_n^r$

e.g. 3!=6 , 5!=120 , 0!=1

$$\binom{7}{3} = \frac{7!}{4!3!} = 35$$

Permutation & Combination

Combinatorial Identity:

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

Analytic proof: (expand both hand sides)

Combinatorial argument: Given n object focus on one of them (obj. 1). There are $\binom{n-1}{r-1}$ groups of size r that contain obj. 1 (since each group contains r-1 other elements out of n-1). Also, there are $\binom{n-1}{r}$ groups of size r, that do not contain obj1. But the total of all r-size groups of n-objects is $\binom{n}{r}$!

Permutation & Combination

Combinatorial Identity:

$$\binom{n}{r} = \binom{n}{n-r}$$

Analytic proof: (expand both hand sides)

Combinatorial argument: Given n objects the number of combinations of choosing any r of them is equivalent to choosing the remaining n-r of them (order-of-objs-not-important!)

Examples

1. Suppose car plates are 7-digit, like AB1234. If all the letters can be used in the first 2 places, and all numbers can be used in the last 4, how many different plates can be made? How many plates are there with no repeating digits?

Solution: a) 26·26·10·10·10·10

b) $P_{26}^{2} \cdot P_{10}^{3} = 26 \cdot 25 \cdot 10 \cdot 9 \cdot 8 \cdot 7$

Examples

2. How many different letter arrangement can be made from the 11 letters of MISSISSIPPI?

Solution: There are: 1 M, 4 I, 4 S, 2 P letters.

Method 1: consider different permutations:

11!/(1!4!4!2!)=34650

Method 2: consider combinations:

$$\binom{11}{1}\binom{10}{4}\binom{6}{4}\binom{2}{2} = \ldots = \binom{11}{2}\binom{9}{4}\binom{5}{4}\binom{1}{1}$$

Examples

3. There are N telephones, and any 2 phones are connected by 1 line. Then how many lines are needed all together?

Solution: $C^2_N = N(N-1)/2$

If, N=5, complete graph with 5 nodes has C^2_5=10 edges.

Examples

4. **N distinct balls** with **M of them white**. Randomly **choose n of the N balls**. What is the probability that the sample contains exactly **m white balls** (suppose every ball is equally likely to be selected)?

Solution: a) For the event to occur, **m out of M white** balls are chosen, and **n-m out of N-M non-white** balls are chosen. And we get $\binom{M}{m}\binom{N-M}{n-m}$

b) Then the probability is $\binom{M}{m}\binom{N-M}{n-m} / \binom{N}{n}$

Later These Probabilities Will be associated with the name HyperGeometric(N, n, M) distrib.

Examples

5. N boys (●) and M girls (○), M<=N+1, stand in 1 line. How many arrangements are there so that no 2 girls stand next to each other?

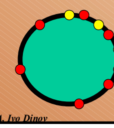
There are **N!** ways of ordering the boys among themselves

There are **M!** ways of ordering the girls among themselves. NOTE – if girls are indistinguishable then there's no need for this factor!

Solution: $N!\cdot\binom{N+1}{M}\cdot M!$

There are **N+1** slots for the girls to fill between the boys And there are **M** girls to position in these slots, hence the coefficient in the middle.

How about they are arranged in a circle?
Answer: $N!\binom{N}{M}M!$

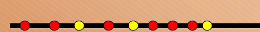


E.g., N=3, M=2

Examples

5a. How would this change if there are N functional (●) and M defective chips (○), M<=N+1, in an assembly line?

Solution: $\binom{N+1}{M}$



There are **N+1** slots for the girls to fill between the boys And there are **M** girls to position in these slots, hence the coefficient in the middle.

Examples

5a. How would this change if there are N functional (●) and M defective chips (○), M<=N+1, in an assembly line?

Solution: $\binom{N+1}{M}$

There are **N+1** slots for the girls to fill between the boys And there are **M** girls to position in these slots, hence the coefficient in the middle.

Binomial theorem & multinomial theorem

Binomial theorem $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$

Deriving from this, we can get such useful formula (a=b=1)

$$\binom{n}{0}+\binom{n}{1}+\dots+\binom{n}{n}=2^n=(1+1)^n$$

Also from $(1+x)^{m+n}=(1+x)^m(1+x)^n$ we obtain:

$$\binom{m+n}{k}=\sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}$$

On the left is the coeff of $1^k x^{(m+n-k)}$. On the right is the same coeff in the product of $(\ldots+ \text{coeff} * x^{(m-i)} +\ldots) * (\ldots+\text{coeff} * x^{(n-k+i)} +\ldots)$.

Multinomial theorem

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$$

Generalization: Divide n distinctive objects into k groups, with the size of every group $n_1,\ldots,n_k$, and $n_1+n_2+\ldots+n_k=n$

$$(x_1+x_2+\ldots+x_k)^n=\sum\binom{n}{n_1,n_2,\ldots,n_k}x_1^{n_1}x_2^{n_2}\ldots x_k^{n_k}$$

where $\binom{n}{n_1,n_2,\ldots,n_k}=\binom{n}{n_1}\binom{n-n_1}{n_2}\ldots\binom{n-n_1-\ldots-n_{k-1}}{n_k}=\frac{n!}{n_1!n_2!\ldots n_k!}$

Multinomial Probabilities $p(n_1,\ldots,n_k)=\frac{n!}{n_1!\cdots n_k!}p_1^{n_1}\cdots p_k^{n_k}$

Multinomial theorem – will discuss in Ch. 03

- N independent trials with results falling in one of k possible categories labeled 1, …, k. Let p_i = the probability of a trial resulting in the i^{th} category, where $p_1+\ldots+p_k=1$
- N_i = number of trials resulting in the i^{th} category, where $N_1+\ldots+N_k = N$
- Ex: Suppose we have 9 people arriving at a meeting.

  P(by Air) = 0.4, *P*(by Bus) = 0.2

  P(by Automobile) = 0.3, *P*(by Train) = 0.1

  P(3 by Air, 3 by Bus, 1 by Auto, 2 by Train) = ?

  P(2 by air) = ?

Examples

7. There are **n balls** randomly positioned in **r distinguishable urns**. Assume n>= r. What is the number of possible combinations? n=9, r=3

1) If the balls are distinguishable (labeled) : r^n possible outcomes, where empty urns are permitted. Since each of the ***n*** balls can be placed in any of the ***r*** urns.

2) If the balls are indistinguishable: **no empty urns** are allowed – select r-1 of all possible n-1 dividing points between the n-balls. $\binom{n-1}{r-1}$

3) If **empty urns** are allowed $\binom{n+r-1}{r-1}$

n=9, r=3, and o are empty bins

Application – Number of integer solutions to linear equ's

1) There are $\binom{n-1}{r-1}$ distinct positive integer-valued vectors $(x_1, x_2 \ldots, x_r)$ satisfying

$$x_1+x_2+\ldots+x_r = n, \quad \& \; x_i>0,\; 1<=i<=r$$

2) There are $\binom{n+r-1}{r-1}$ distinct positive integer-valued vectors $(y_1, y_2 \ldots, y_r)$ satisfying

$$y_1+y_2+\ldots+y_r = n, \quad \& \; y_i>=0,\; 1<=i<=r$$

Since there are *n+r-1* possible positions for the dividing splitters (or by letting $y_i=x_i-1$, RHS=n+r).

Example

1) An investor has $20k to invest in 4 potential stocks. Each investment is in increments of $1k, to minimize transaction fees. In how many different ways can the money be invested?

2) $x_1+x_2+x_3+x_4=20,\; x_k>=0$ ➔ $\binom{23}{3} = 1{,}771$

3) If not all the money needs to be invested, let x5 be the left over money, then $x_1+x_2+x_3+x_4+x_5=20$ $\binom{24}{4} = 10{,}626$

Examples

8. Randomly give n pairs of **distinctive shoes** to n people, with 2 shoes to everyone. How many arrangements can be made? How many arrangements are there, so that everyone gets an original pair? What is the the probability of the latter event, ***E***?

Solution: a) according to $\binom{2n}{n_1, n_2, \ldots, n_r} =$

Note: **r = n = # of pairs**!

total arrangements is $\binom{2n}{n_1}\binom{2n-n_1}{n_2}\ldots\binom{2n-n_1-n_2-\ldots-n_{r-1}}{n_r} =$

$N=(2n)!/(2!)^r = (2n)!/2^r$ $\quad \frac{(2n)!}{n_1!n_2!\ldots n_r!} = \frac{(2n)!}{(2!)^r}$

b) Regard every shoe pair as one object, and give them to people, there are M=n! arrangements.

c) $P(\boldsymbol{E})=M/N= n!/[(2n)!/2^n]=1/(2n-1)!!$ (Do n=6, case by hand!)

*note: $n!!=n(n-2)(n-4)\ldots$

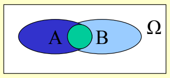
Sterling Formula for asymptotic behavior of n!

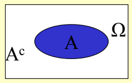
Sterling formula:

$$n! = \sqrt{\frac{2\pi}{n}} \times \left(\frac{n}{e}\right)^n$$

Probability and Venn diagrams

Venn's diagram





Union: A∪B

Intersection: A∩B

A^c denotes the part in Ω but not in A.

Properties:

$A\cap B=B\cap A$,	$A\cup B=B\cup A$,
$(A\cap B)\cap C=A\cap(B\cap C)$,	$(A\cup B)\cup C=A\cup(B\cup C)$,
$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$	$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$

De Morgan's Law: $A^c\cap B^c=(A\cup B)^c$, $A^c\cup B^c=(A\cap B)^c$

Generalized: $(\cap E_i)^c=\cup E_i^c$, $(\cup E_i)^c=\cap E_i^c$, $i=1,2,\ldots,n$

Probability and Venn diagrams

Proposition

$P(A_1\cup A_2\cup\ldots\cup A_n)=$

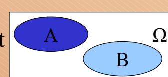
$$\sum_{i=1}^{n} P(A_i)-\sum_{1\le i1<i2\le n} P\left(A_{i1}\bigcap A_{i2}\right)+\ldots$$

$$+(-1)^{r+1}\sum_{1\le i1<j2<\ldots<ir\le n} P\left(A_{i1}\bigcap A_{i2}\bigcap \ldots\bigcap A_{ir}\right)+\ldots$$

$$+(-1)^{n+1} P\left(A_{i1}\bigcap A_{i2}\bigcap \ldots\bigcap A_{in}\right)$$

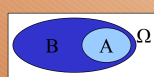
Probability and Venn diagrams

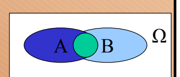
Exclusive events:statistically independent



$A\cap B=\Phi$ or $P(A\cap B)=0$

Conditional probability:





P(A | B)=P($A\cap B$)/P(B)

$A=AB\cup AB^C$ or ***$P(A)=P(A|B)+P(A|B^C)$***

Examples

9. (True or false)All K are S, all S not W. Then all W are not K. (T)
All K are S, Some S are W. Then surely some K is W. (F)

10. A class have 100 pupils, each of them is enrolled in at least one course among A,B&C. It is known that 35 have A, 40 have B,50 have C, 8 have both A&B,12 have both A&C, 10 have both B&C. How many pupils have all 3 courses?

Solution: Use Venn's diagram,35+40+50-8-12-10+X=100

→ X = 5

Note: The arrangement: 8→A&B; 15→A&C; 12→B&C won't work, since the only solution is 10→ A&B&C, but A&B&C<=A&B, which is a contradiction!