ECO239 Statistics I

Week 6

Remaining from last week

Probability

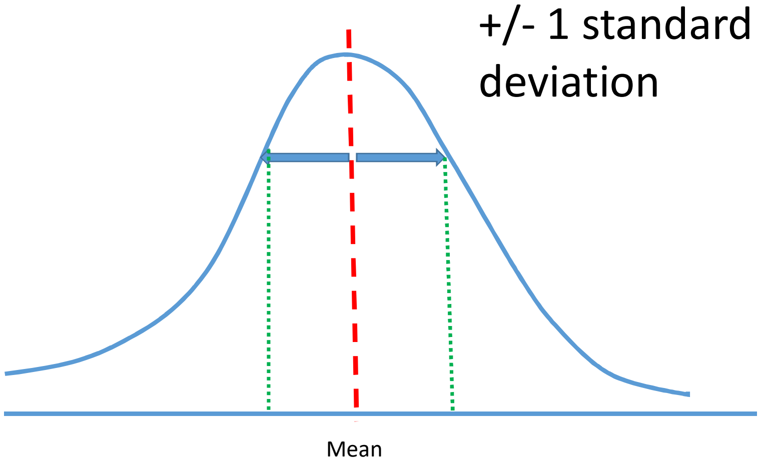
Chebyshev's Theorem

- Answers the question "How much percentage of observations can be found in the interval

$$\mu \pm k\sigma \quad ?"$$

More generally…

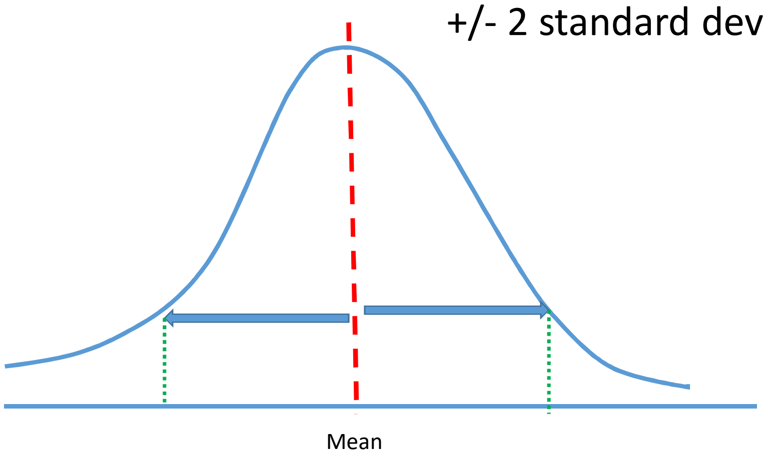
+/- 1 standard deviation



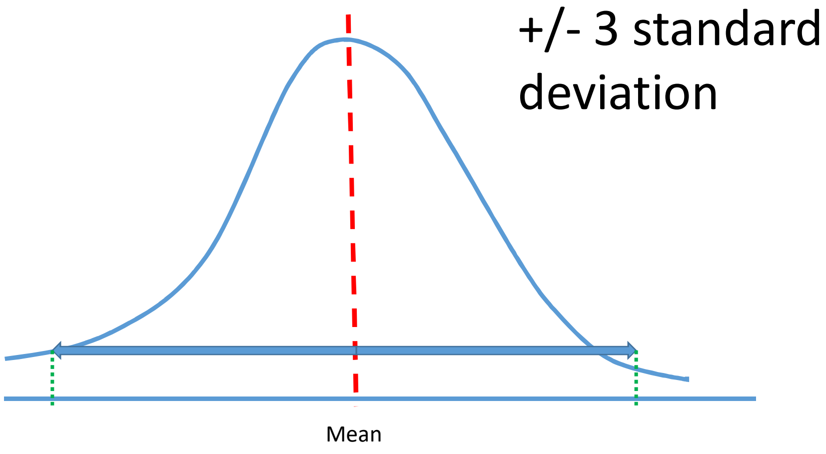
Chevyshev's theorem does not work for this case, k = 1.

More generally…

+/- 2 standard deviation



More generally…



Chebyshev's Theorem

- For any mean and standard deviation, and **k >1**, the % of observation that fall within the interval $\mu \pm k\sigma$ is at least

$$100\left[1-\left(\frac{1}{k^2}\right)\right]\%$$

- Regardless of how the data are distributed.
- Does not work for k = 1.

Q: for k=2, what is the % of observations? How about for k = 3?

Within	At least
K=2 (mean +/- 2 stdev)	(1-(1/(2^2))*100% = 75%
K=3 (mean +/- 3 stdev)	(1-(1/(3^2))*100% = 89%

- Does not work for k = 1.
- K does not have to be integers.

$$100\left[1-\left(\frac{1}{k^2}\right)\right]\%$$

Chebyshev's Theorem

- **Advantage**: Applicable to any population & distributional shapes.
- **Disadvantage**: In reality, distributions are relatively close to symmetric, and % of observations in a specific range is much higher.

Practice

$$100\left[1-\left(\frac{1}{k^2}\right)\right]\%$$

- A large class with 280 students.
- Midterm exam result: mean = 74, stdev=6.
- At least how many students scored between 50 and 98 according to Chevyshev's Theorem?

If stdev = 8, instead of 6,

- At least how much % of students are included in the same range (50 & 98) ?
- Do you think it's more /less than the previous question? And WHY?

Practice

- In company A, the average salary is 6000 TL with standard deviation of 1200 TL.
- According to Chevyshev's theorem, what is the interval in which at least 80% of the salaries lie?

$$100\left[1-\left(\frac{1}{k^2}\right)\right]\%$$

$$\mu \pm k\sigma$$

Empirical Rule

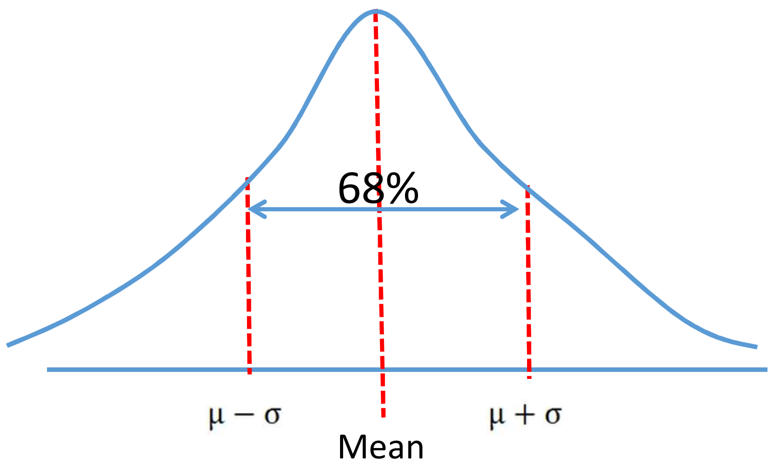
- If the data distribution is bell-shaped,

$\mu \pm \sigma$ contains about 68% of observations

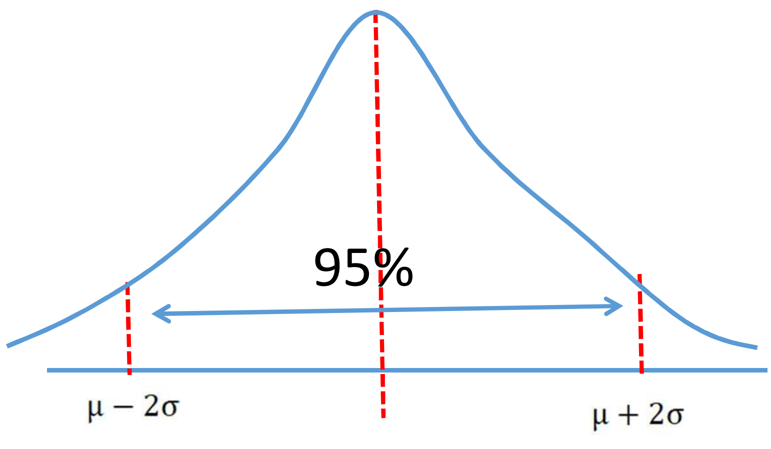
$\mu \pm 2\sigma$ contains about 95% of observations

$\mu \pm 3\sigma$ contains about 99.7% of observations

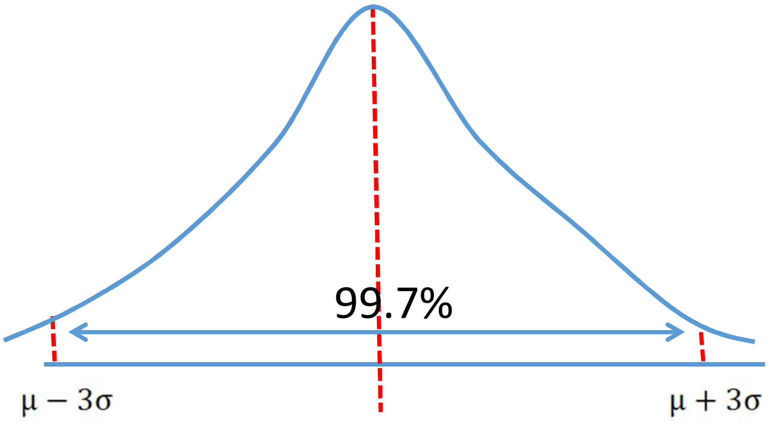
$\mu \pm \sigma$ contains about 68% of observations



$\mu \pm 2\sigma$ contains about 95% of observations



$\mu \pm 3\sigma$ contains 99.7% of observations



Practice

- n = 280 students
- μ = 74, σ = 6.

Q1: How many students scored between 62 and 86 according to Empirical Rule?

Q2: If you score 92, you are in top ______ %.

Practice

- n = 280 students
- μ= 74, σ = 6.

Q1: How many students scored between 62 and 86 according to Empirical Rule?

Q2: If you score 92, you are in top ______ %.

Empirical vs. Chevyshev

- Compare this result with Chevyshev's theorem.
- n = 280 students; mean= 74, stdev= 6.

Q1: How many students scored between 62 and 86 according to Empirical Rule?

Q1': How many students scored between 62 and 86 according to Chevyshev's theorem?

Practice

- Average bill in Quick China, mean = 55 TL,st.dev. = 8.3 TL.

Q: 99.7% of the time you expect your bill to be between [ ] and [ ] TL according to Empirical Rule

Q:95% of the time you expect your bill to be between [ ] and [ ] TL according to Empirical Rule.

Q: According to Chevyshev's Theorem, 95% will be included within [,] range.

Introduction to Probability

Important terms

1. **Random Experiment** – a process leading to an uncertain outcome

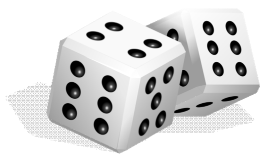
e.g. coin toss

die rolling

2. **Basic Outcomes-** a possible outcome of a random experiment

e.g. coin toss: Head or Tail

die rolling: 1, 2, 3, 4, 5, 6



3. **Sample space** – the collection of all possible outcomes of a random experiment.

e.g. coin toss S = [H, T]

die rolling S = [1,2,3,4,5,6]

4. **Event**- any subset of basic outcomes from the sample space.

e.g. die rolling: odd number { 1, 3, 5} ,even number {2, 4, 6}, less than 3 {1, 2}…

*Null event = absence of a basic outcome.

$\emptyset$

5. **Intersection of Event**

If A and B are two events in a sample space S, then the intersection, $A \cap B$, is the set of all outcomes in S that belong to both A and B.

e.g. event A: {1, 3, 5}, event B: {1, 2, 3}

=> $A \cap B = \{1, 3\}$

6. **Mutually Exclusive Events**

Events A and B are mutually exclusive events if they have no basic outcomes in common.

$A \cap B = \emptyset$

e.g. A={1, 3, 5} , B={2, 4, 6}

$A \cap B = \emptyset$ => A and B are mutually exclusive.

7. **Union** of Events

Union A U B is the set of all outcomes in S that belong to either A or B.

A={1, 3, 5}

B={3, 4, 5}

AUB={1, 3, 4, 5}

8. Events E1, E2,… Ek are Collectively Exhaustive events if E1 U E2 U…U Ek = S.

(Union of all events = sample space itself)

e.g. Rolling a die case.

E1 = {1, 2}, E2 = {3, 4}, E3 = {5, 6}.

E1 U E2 U E3 = S.

e.g. Rolling a die case.

E1= {1, 3, 5}, E2 = {2, 4, 6}, E3 = {3, 4, 5}

E1 U E2 U E3 = S.

9. Complement of an event A is the set of all basic outcome in the sample space that do not belong to A.

$$A \cup \bar{A} = S$$

e.g. Rolling a die

A = {1, 3, 5}

$\bar{A}$= {2, 4, 6}

Practice

- Rolling a die
- A= {2, 3, 6}, B = {4, 5, 6}

$A \cap B =$

AUB=

$A \cup \bar{B}$

$\bar{A} \cup B$

$\bar{A} \cap B$

$A \cap \bar{B}$

Are A and B Mutually Exclusive?

Are A and B Collectively Exhaustive?

Important Result

1. $$(A \cap B) \cup (\bar{A} \cap B) = B$$

2. $$A \cup (\bar{A} \cap B) = A \cup B$$

3. E1, E2, … Ek: k mutually exclusive and collectively exhaustive events.

"A" is some other event.

$$(E_1 \cap A) \cup (E_2 \cap A) \cup \ldots \cup (E_k \cap A) = A$$

Probability

- The chance that an uncertain event will occur.
- P(A) = Probability of an event A
- P(A) = N_A/N = (# of outcomes that satisfying the event) / (total # of outcome in the sample space).
- $0 \leq P(A) \leq 1$

Practice

Which of the following events would you

be most surprised by?

(a) exactly 3 heads in 10 coin flips

(b) exactly 3 heads in 100 coin flips

(c) exactly 3 heads in 1000 coin flips

Practice

Which of the following events would you

be most surprised by?

(a) exactly 3 heads in 10 coin flips

P(A) = 3/10 = 0.3.

(b) exactly 3 heads in 100 coin flips

P(A)=3/100=0.03.

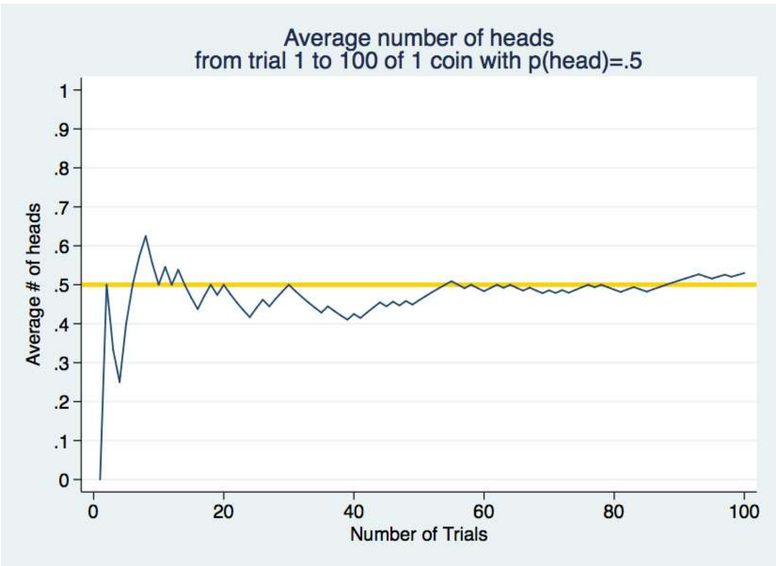
(c) exactly 3 heads in 1000 coin flips

P(A)=3/1000=0.003.

Law of large numbers

Law of large numbers states that as more observations are collected, the proportion of occurrences with a particular outcome, $\hat{p}n$, converges to the probability of that outcome, p.

LLN: Example "Coin Toss"



Let's Try

	H or T		H or T		H or T
1		11		21	
2		12		22	
3		13		23	
4		14		24	
5		15		25	
6		16		26	
7		17		27	
8		18		28	
9		19		29	
10		20		30	

Law of large numbers (cont.)

When tossing a fair coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss?

(a) 0.5

(b) less than 0.5

(c) more than 0.5

H H H H H H H H H H ?

Law of large numbers (cont.)

When tossing a fair coin, if heads comes up on each of the first 10 tosses, what do you think the chance is that another head will come up on the next toss? 0.5, less than 0.5, or more than 0.5?

$\underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{H}\ \underline{?}$

- The probability is still 0.5, or there is still a 50% chance that another head will come up on the next toss.

$P(H \text{ on } 11^{th} \text{ toss}) = P(T \text{ on } 11^{th} \text{ toss}) = 0.5$

Practice

What is the probability of drawing **a jack or a red card** from a well shuffled full deck?

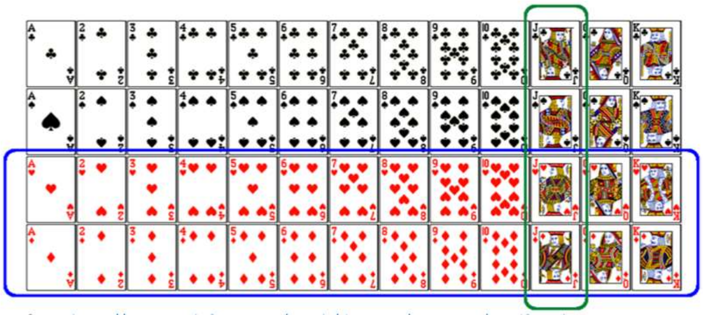
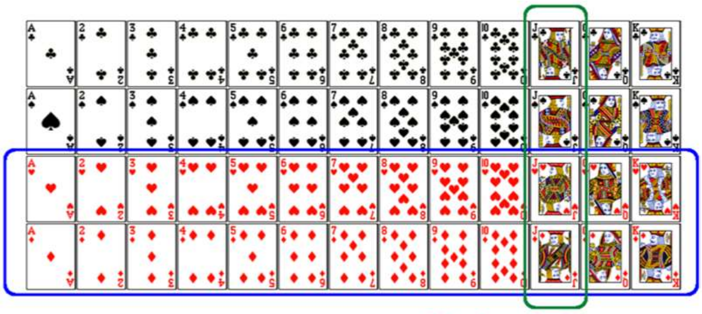


Figure from http://www.milefoot.com/math/discrete/counting/cardfreq.htm

P(AUB)=P(A)+P(B)-P(A∩B)

What is the probability of drawing a jack or a red card from a well shuffled full deck?



$$P(\text{jack or red}) = P(\text{jack}) + P(\text{red}) - P(\text{jack and red})$$
$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52}$$

Figure from http://www.milefoot.com/math/discrete/counting/cardfreq.htm

Practice

What is the probability that a randomly sampled student thinks marijuana should be legalized or they agree with their parents' political views?

	Share Parents' Politics		
Legalize MJ	No	Yes	Total
No	11	40	51
Yes	36	78	114
Total	47	118	165

(a) (40 + 36 - 78) / 165

(b) (114 + 118 - 78) / 165

(c) 78 / 165

(d) 78 / 188

(e) 11 / 47

Recap

General addition rule

P(A or B) = P(A) + P(B) - P(A and B)

Note: For mutually exclusive (disjoint) events
P(A and B) = 0, so the above formula simplifies to
P(A or B) = P(A) + P(B)

Practice

- 75% of customers use mustard
- 80% of customers use ketchup
- 65% of customers use both.

Q: What is the probability that a customer uses at least one?

Q: What is the probability that a customer uses none of them?

Practice

- 75% of customers use mustard
- 80% of customers use ketchup
- 65% of customers use both.

Q: What is the probability that a customer uses at least one?

Q: What is the probability that a customer uses none of them?

Combination Formula

- # of combination to pick K out of n

$$C_k^n = \frac{n!}{k!\,(n-k)!}$$

Practice $C_k^n = \dfrac{n!}{k!\,(n-k)!}$

- Ayse will choose 3 laptops out of 20. What is the probability of choosing 2 HP and 1 SONY out of 10 HP, 5 SONY and 5 DELL?

Probability distributions

A *probability distribution* lists all possible events and the probabilities with which they occur.

- The probability distribution for the gender of one kid:

Event	Male	Female
Probability	0.5	0.5

Probability distributions

A *probability distribution* lists all possible events and the probabilities with which they occur.

- The probability distribution for the gender of one kid:

Event	Male	Female
Probability	0.5	0.5

- Rules for probability distributions:
  1. The events listed must be disjoint
  2. Each probability must be between 0 and 1
  3. The probabilities must total 1

Probability distributions

A *probability distribution* lists all possible events and the probabilities with which they occur.

- The probability distribution for the gender of one kid:

Event	Male	Female
Probability	0.5	0.5

- Rules for probability distributions:
  1. The events listed must be disjoint
  2. Each probability must be between 0 and 1
  3. The probabilities must total 1
- The probability distribution for the genders of two kids:

Event	*MM*	*FF*	*MF*	*FM*
Probability	*0.25*	*0.25*	*0.25*	*0.25*

Conditional Probability

Relapse

Researchers randomly assigned 72 chronic users of cocaine into three groups: desipramine (antidepressant), lithium (standard treatment for cocaine) and placebo. Results of the study are summarized below.

	relapse	no relapse	total
desipramine	10	14	24
lithium	18	6	24
placebo	20	4	24
total	48	24	72

http://www.oswego.edu/~srp/stats/2_way_tbl_1.htm

Marginal probability

What is the probability that a patient relapsed?

	relapse	no relapse	total
desipramine	10	14	24
lithium	18	6	24
placebo	20	4	24
total	48	24	72

Marginal probability

What is the probability that a patient relapsed?

	relapse	no relapse	total
desipramine	10	14	24
lithium	18	6	24
placebo	20	4	24
total	48	24	72

P(relapsed) = 48 / 72 ~ 0.67

Joint probability

What is the probability that a patient received the antidepressant (desipramine) and relapsed?

	relapse	no relapse	total
desipramine	10	14	24
lithium	18	6	24
placebo	20	4	24
total	48	24	72

Joint probability

What is the probability that a patient received the antidepressant (desipramine) and relapsed?

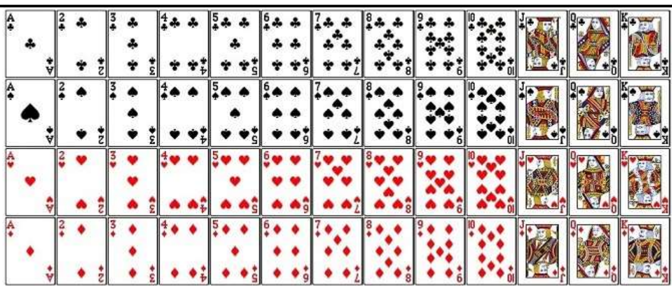
	relapse	no relapse	total
desipramine	10	14	24
lithium	18	6	24
placebo	20	4	24
total	48	24	72

P(relapsed and desipramine) = 10 / 72 ~ 0.14

Probability Table

	A	A_complement $\bar{A}$	
B	$A \cap B$	$\bar{A} \cap B$	$P(B)$
B_completment $\bar{B}$	$A \cap \bar{B}$	$\bar{A} \cap \bar{B}$	$P(\bar{B})$
	$P(A)$	$P(\bar{A})$	1

Q: Complete the Probability Table



- Let event A = cards is an Ace {H1, D1, C1, S1}
- Let event B = cards is a red suit {H1~13, D1~13}

	A	A_complement $\bar{A}$	
B	$A \cap B$	$\bar{A} \cap B$	$P(B)$
B_completment $\bar{B}$	$A \cap \bar{B}$	$\bar{A} \cap \bar{B}$	$P(\bar{B})$
	$P(A)$	$P(\bar{A})$	1