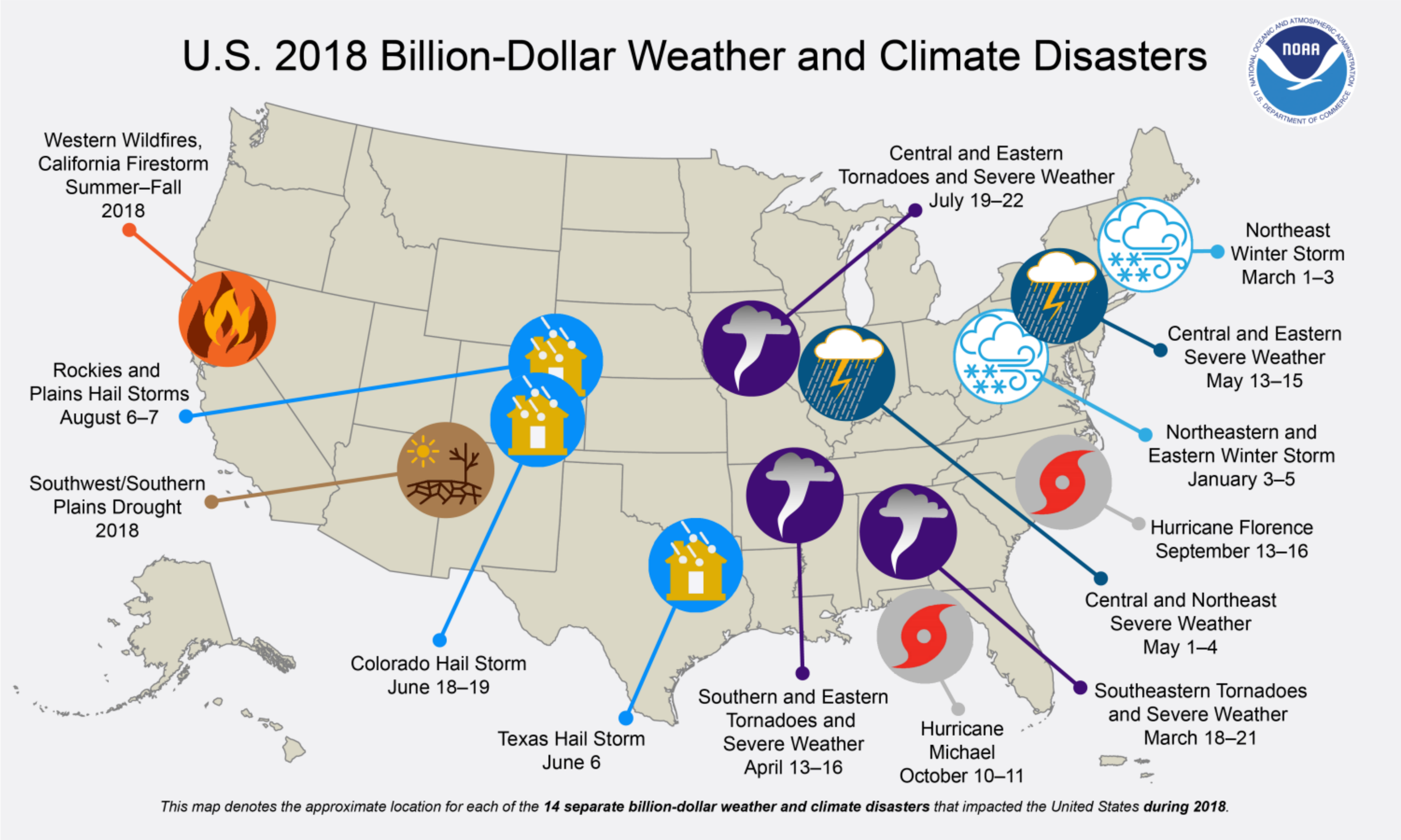
2018 Billion Dollar Weather Disasters



What about all the major weather events and weather-related fatalities in 2018?

Storm Events Database

The Storm Events Database contains the records used to create the official NOAA Storm Data publication, documenting:

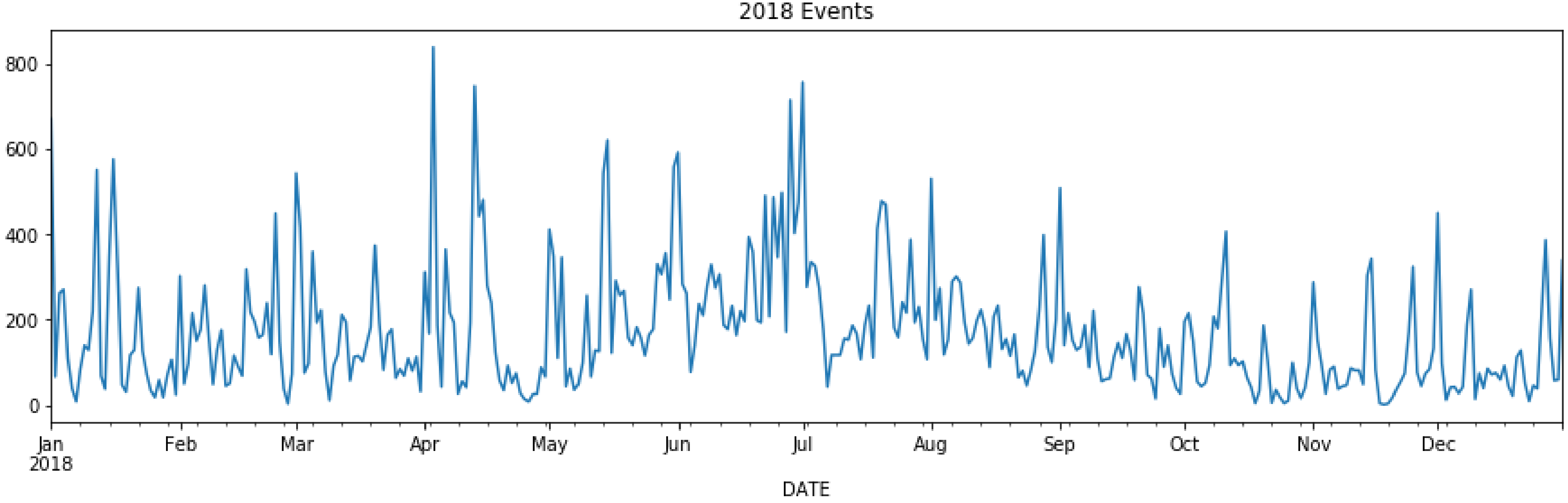
a. The occurrence of storms and other significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, and/or disruption to commerce;

b. Rare, unusual, weather phenomena that generate media attention, such as snow flurries in South Florida or the San Diego coastal area; and

c. Other significant meteorological events, such as record maximum or minimum temperatures or precipitation that occur in connection with another event.

The database currently contains data from **January 1950 to July 2019**, as entered by NOAA's National Weather Service (NWS). Due to changes in the data collection and processing procedures over time, there are unique periods of record available depending on the event type. NCEI has performed data reformatting and standardization of event types but has not changed any data values for locations, fatalities, injuries, damage, narratives and any other event specific information. Please refer to the Database Details page for more information.

In 2018
62351 Weather Events
799 fatalities
>$35 billion in property and crop losses



Probability Definitions

- Event- possible uncertain outcomes
- Null event- can’t happen
- Elementary event- can’t be decomposed into other events
- Compound event- decomposable into 2 or more elementary events
- S- sample or event space- all possible elementary events
- Mutually exclusive- two events that can’t occur at same time
- MECE- Mutually exclusive and collectively exhaustive- no more than 1 event can occur and at least one event will occur

More Definitions

- E- Event
- Pr{E}- probability of Event E; $0 \leq Pr\{E\} \leq 1$
- Pr{E} = 0, event does not occur
- Pr{E} = 1 event occurs

Temperature below Precipitation below	Temperature above Precipitation below
Temperature below Precipitation above	Temperature above Precipitation above

Figure 4.2. MECE possibilities for seasonal forecasts temperature and precpitation anomalies for a specifi location.

Two Statistical Frameworks: Frequency vs. Bayesian

- Frequency- probability of an event is its relative frequency after many trials
- a- number of occurrences of E
- n- number of opportunities for E to take place
- a/n- relative frequency of E occurring
- $Pr\{E\} \rightarrow a/n$ as $n \rightarrow \infty$

Two Statistical Frameworks: Frequency vs. Bayesian

- Bayesian- probability represents the degree of belief of an individual about an outcome of an uncertain event
- Some events occur so rarely that there is no long-term relevant probability
- Two individuals can have different probabilities for same outcome
- Bookies are Bayesian
  - **Super Bowl odds 2019: Heavy Patriots money had sportsbooks rooting for Rams**

Deck of Cards

- If don’t look at cards, then the odds of getting any specific card when dealt two cards is the same as if you only were dealt one card: they are independent events
- The following table applies to dealing two independent cards to you. The probability of getting one ace on either card is 4/52=7.7%

Pr{ace}	**Pr{10-K}**	**Pr{2-9}**	**Pr{21}**
7.7	30.8	61.5	4.8

MECE

First Card

Second card	2-9	10-K	Ace	totals
2-9	37.9%	18.9%	4.7	61.5%
10-K	18.9%	9.5%	2.4%	30.8%
Ace	4.7	2.4%	.6%	7.7%
totals	61.5%	30.8%	7.7%	100%

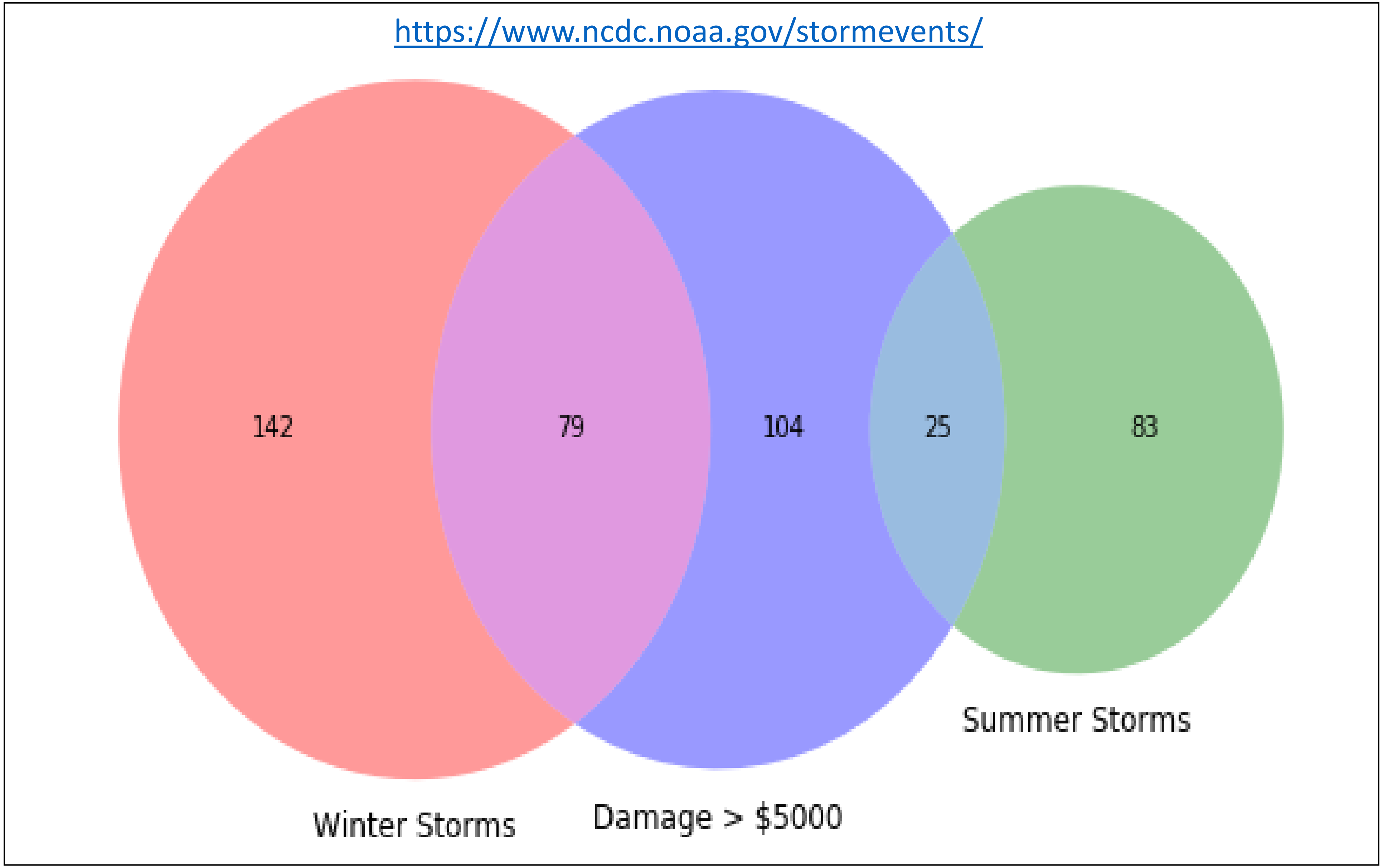
The odds of getting both cards to be aces is 4/52 * 4/52 = 0.6%

Deck of Cards

- If don’t look at cards, then the odds of getting any specific card when dealt two cards is the same as if you only were dealt one card: they are independent events
- The following table applies to dealing two independent cards to you. The probability of getting one ace on either card is 4/52=7.7%

Pr{ace}	**Pr{10-K}**	**Pr{2-9}**	**Pr{21}**
7.7	30.8	61.5	4.8

	Tally of occurrences	**Total number of occurrences (a)**	**Total # of opportunities (n)**	**Observed probability a/n *100 (%)**	**Expected Probability from Part 1 (%)**	**Expected number of outcomes if n opportunities**
ace			20		7.7	1.5
10-K			20		30.8	6
2-9			20		61.5	12
21			10		4.8	.5



Number of Opportunities: 2340 (180 days * 13 years)

More concepts

- $\{E\}^c$- complement of $\{E\}$, event does not occur
- $Pr\{E\}^c = 1-\{E\}$
- $Pr\{E_1 \cap E_2\}$- joint probability that E_1 & E_2 occur
- $Pr\{E_1 \cap E_2\} = 0$ if E_1 & E_2 are mutually exclusive
- $Pr\{E_1 \text{ U } E_2\}$- probability that E_1 OR E_2 occur
- $Pr\{E_1 \text{ U } E_2\} = Pr\{E_1\} + Pr\{E_2\} - Pr\{E_1 \cap E_2\}$

Simplified #'s from those in Chapter 3 Notes.

Total number of opportunities = 1000.

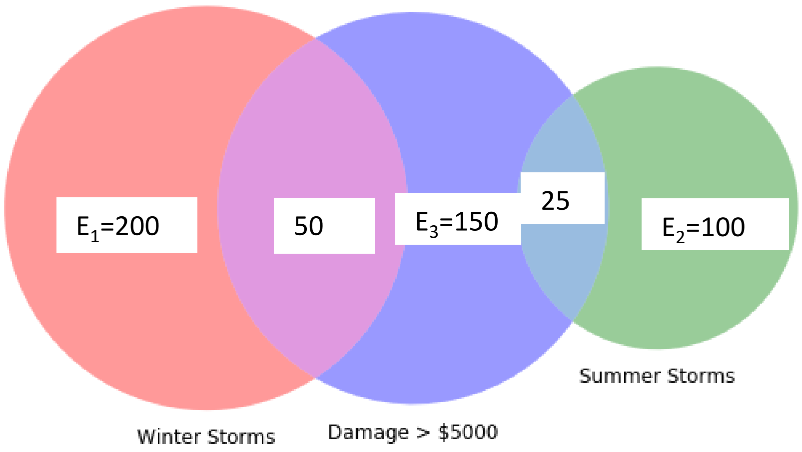
$\{E_1\}$ = occurrence of winter storms = 200.

$\{E_2\}$ = occurrence of summer storms = 100.

$\{E_3\}$ = occurrence of property damage = 150.

$\{E_1 \cap E_2\} = 0$.

$\{E_1 \cap E_3\} = 50$.



$Pr\{E_1\}$	20%	$Pr\{E_3 \mid E_1\}$	
$Pr\{E_2\}$		$Pr\{E_3 \mid E_2\}$	
$Pr\{E_3\}$		$Pr\{E_1 \mid E_3\}$	
$Pr\{E_1 \cap E_2\}$		$Pr\{E_1 \cup E_2 \cup E_3\}$	
$Pr\{E_2 \cap E_3\}$		$Pr\{E_1 \cup E_2 \cup E_3\}^c$	
$Pr\{E_1 \cap E_3\}$		$Pr\{E_1\}^c$	
$Pr\{E_1 \cup E_2$		$Pr\{E_1 \cup E_2\}^c$	
$Pr\{E_2 \cup E_3\}$		$Pr\{E_1 \cap E_2\}^c$	
$Pr\{E_1 \cup E_3\}$			

Conditional Probability

- Conditional probability: probability of $\{E_2\}$ given that $\{E_1\}$ has occurred
- $Pr\{E_2 \mid E_1\} = Pr\{E_1 \cap E_2\} / Pr\{E_1\}$
- E_1 is the conditioning event
- If E_1 and E_2 are independent of each other, then $Pr\{E_2 \mid E_1\} = Pr\{E_2\}$ and $Pr\{E_1 \mid E_2\} = Pr\{E_1\}$
- Fair coin- Pr{heads} = 0.5
  - chance of getting heads on second toss is independent of the first

    Pr{heads | heads} = 0.5

    Pr{heads} twice = 0.5 * 0.5 = .25

Bayes Theorem

- $\Pr\{E_2 \mid E_1\} = \Pr\{E_1 \cap E_2\} / \Pr\{E_1\}$
- E_1 is the conditioning event
- What is the advantage? Probability of conditioning event E_1 only computed once

- $\Pr\{E_1 \mid E_2\} = \Pr\{E_2 \mid E_1\} * \Pr\{E_1\} / \Pr\{E_2\}$

- $\Pr\{E_1 \cap E_2\} = \Pr\{E_2 \mid E_1\} * \Pr\{E_1\}$
- $\Pr\{E_1 \cap E_2\} = \Pr\{E_1 \mid E_2\} * \Pr\{E_2\}$

Bayesian Application: how you should respond to “evidence”

	Pos Test	Neg Test	TOTAL
DRUG USER	0.495%	0.005%	0.5%
NOT DRUG USER	.995%	98.505%	99.5%
TOTAL	1.49%	98.51%	1

How many drug users test positive? 99%
*Out of a ten thousand people, how many drug users do they catch? .00495*10000=~50*

What are odds of falsely accusing non drug user?
E_1 – not drug user
E_2 - positive test
$Pr\{E_1\}$ – 99.5%
$Pr\{E_2\}$ – 1.49%
$Pr\{E_2 \mid E_1\}$ – .995%
$Pr\{E_1 \mid E_2\} = Pr\{E_2 \mid E_1\} * Pr\{E_1\} / Pr\{E_2\} = 0.995 * 99.5 / 1.49 = 68\%$

- If you are a non drug user, you have a 68% chance of getting a “false positive”
- *How many non-drug users get falsely accused?.00995*10000=~100*

Conclusion: always always ask for second opinion if clean and test positive

Application of Bayes theorem: how to be rational responding to probabilities

	Pos Test	Neg Test	TOTAL
DRUG USER	0.495%	0.005%	0.5%
NOT DRUG USER	.995%	98.505%	99.5%
TOTAL	1.49%	98.51%	1

What are odds of a drug user skating?

E_4 – drug user

E_3 - negative test

$Pr\{E_3\}$ – 98.51%

$Pr\{E_4\}$ – 0.5%

$Pr\{E_3 \cap E_4\}$ – .005%

$Pr\{E_4 \mid E_3\} = Pr\{E_4 \cap E_3\} / Pr\{E_3\} =$

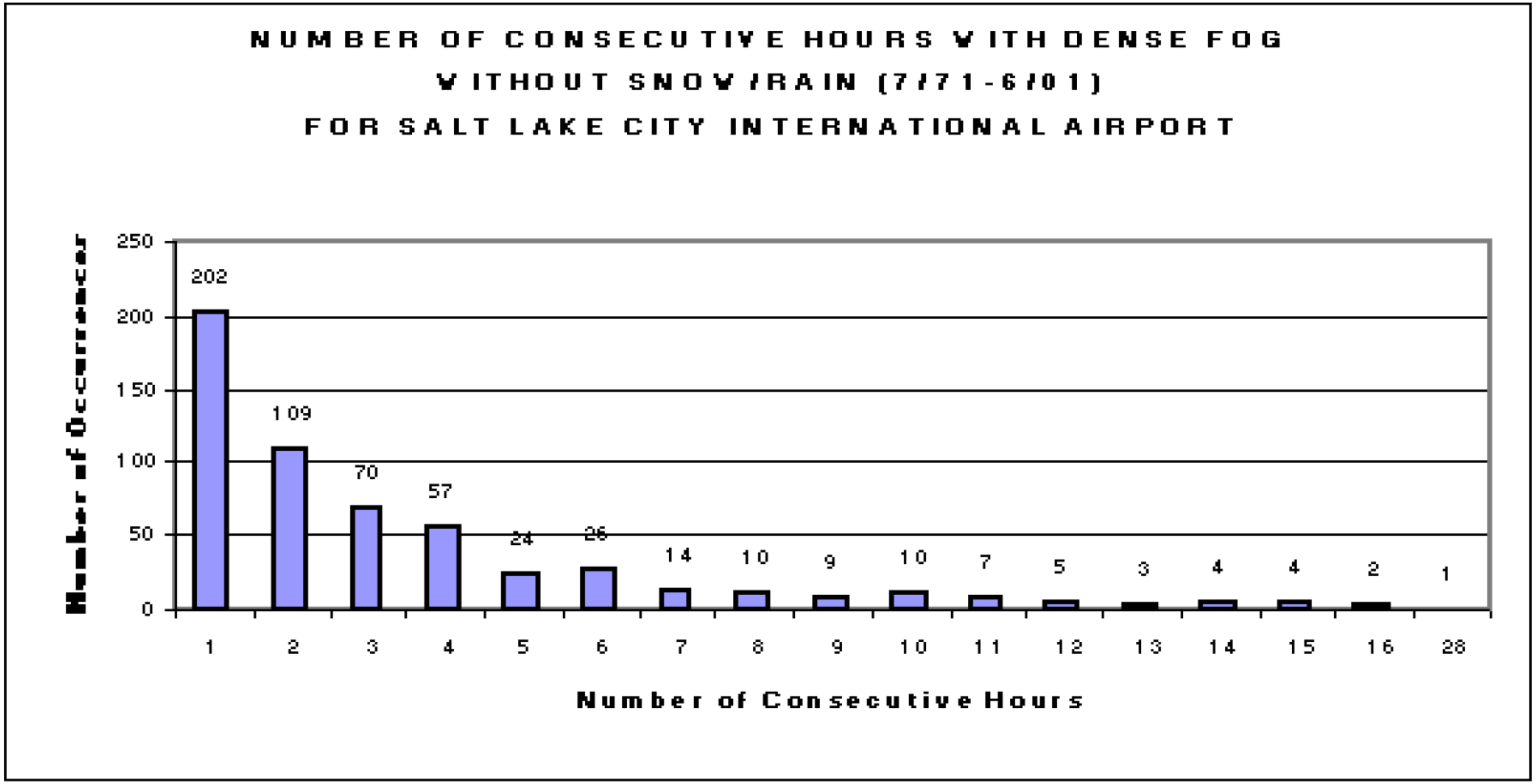
$= 0.005 / 98.51 = .0051\%$

Out of 10000 people, maybe 1 drug user will test negative

Conclusion: people who give drug tests are more interested in making sure drug users are caught than worrying about innocent people being falsely accussed

Olympic Fog Climatology

When phenomena are persistent, then “odds” are higher once an event is underway



Forecast Verification

- What is your reason for doing it?
- (Brier and Allen 1951; Compendium of Meteorology)
  - Administrative: who's blowing the forecasts?
  - Scientific: why do errors happen?
  - Economic: what's the impact of forecast errors?

Measures oriented: “give me a number!”

- Distill set of forecasts and observations into small # of metrics

		Observed	Observed	Forecast marginal totals
		Yes	No	
Forecast	Yes	a	b	a+b
Forecast	No	c	d	c+d
	Observed marginal totals	a+c	b+d	n=a+b+c+d sample size

- PC = percent correct = $\frac{a+d}{n}$
- FAR = false alarm ratio = $\frac{b}{a+b}$
- TS = CSI = $\frac{a}{a+b+c}$
- POD = HR = $\frac{a}{a+c}$

What if it just happened by chance?

		Observed	Observed	Forecast marginal totals
		Yes	No	
Forecast	Yes	a	b	a+b
Forecast	No	c	d	c+d
	Observed marginal totals	a+c	b+d	n=a+b+c+d sample size

- Random correct yes forecast by chance = $\frac{(a+b)}{n}\frac{(a+c)}{n}$
- Random correct no forecast by chance = $\frac{(b+d)}{n}\frac{(c+d)}{n}$
- $SS = \frac{(correct\ forecasts - random\ correct\ forecasts)}{(total\ forecasts - random\ correct\ forecasts)}$
- $HSS = \frac{2(ad-bc)}{(a+c)(b+d)+(a+b)(b+d)}$

Verifying wind forecasts

		Observed	Observed	Forecast Marginal totals
		≥ 5m/s	<5 m/s	
Forecast	≥ 5m/s	11	6	17
Forecast	<5 m/s	16	44	60
	Observed Marginal totals	27	50	77

PC= 71.4%; FAR= 35.3%; TS= 33.3%; and POD = 40.7%
randomly correct yes forecast: 7.7%
randomly correct no forecast: 50.1%
HSS= 31.4%

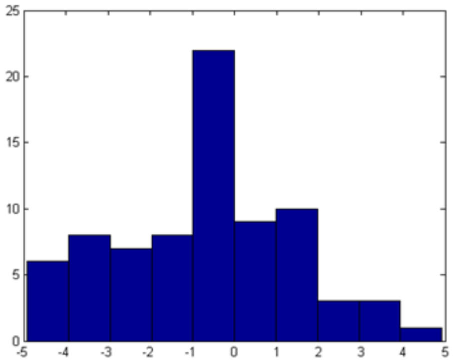
Distributions oriented: “ how close am I?”

- Assessing the characteristics of joint distribution of errors
- Categorize errors: which errors are smallest, which are biggest as a function of the range of values?
- Relies heavily on conditional probabilities
- http://meso1.chpc.utah.edu/jfsp/

Forecast Verification

- http://meso1.chpc.utah.edu/jfsp/
- Select Wildfires by WFO
- Select SLC
- Look at over all years, then focus on wildfires in Utah in 2016
- Then follow along in class

Assessing Forecast Accuracy



		Observed	Observed	Observed	Error Marginal totals
		≤3 m/s	3-6 m/s	≥6 m/s	
Error	≤ -2 m/s	0	10	11	21
Error	± 2 m/s	22	20	7	49
Error	> 2 m/s	0	7	0	7
	Observed Marginal totals	22	37	18	77

- 26% of the forecasts were within 2 m/s when the wind speeds were between 3 and 6 m/s (20/77)
- Given that the observed wind speed is greater than 6 m/s: ($Pr\{E_1\}$ = 18/77= 23.4%)
- Probability that the forecasters predict strong winds to be too light $Pr\{E_2 | E_1\}$:

$Pr\{E_2 | E_1\} = Pr\{E_1 \cap E_2\} / Pr\{E_1\} = ((11/77)/(18/77)) = 64.7\%$

What you should be doing

- Read all of Chapter 3 Notes
- Assignment 5 Extra Credit . Due Nov. 15
- Turn in today’s in class assignment
- You will be doing an exploratory data assignment in class on Thursday. It will be due after next Tuesday.