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:

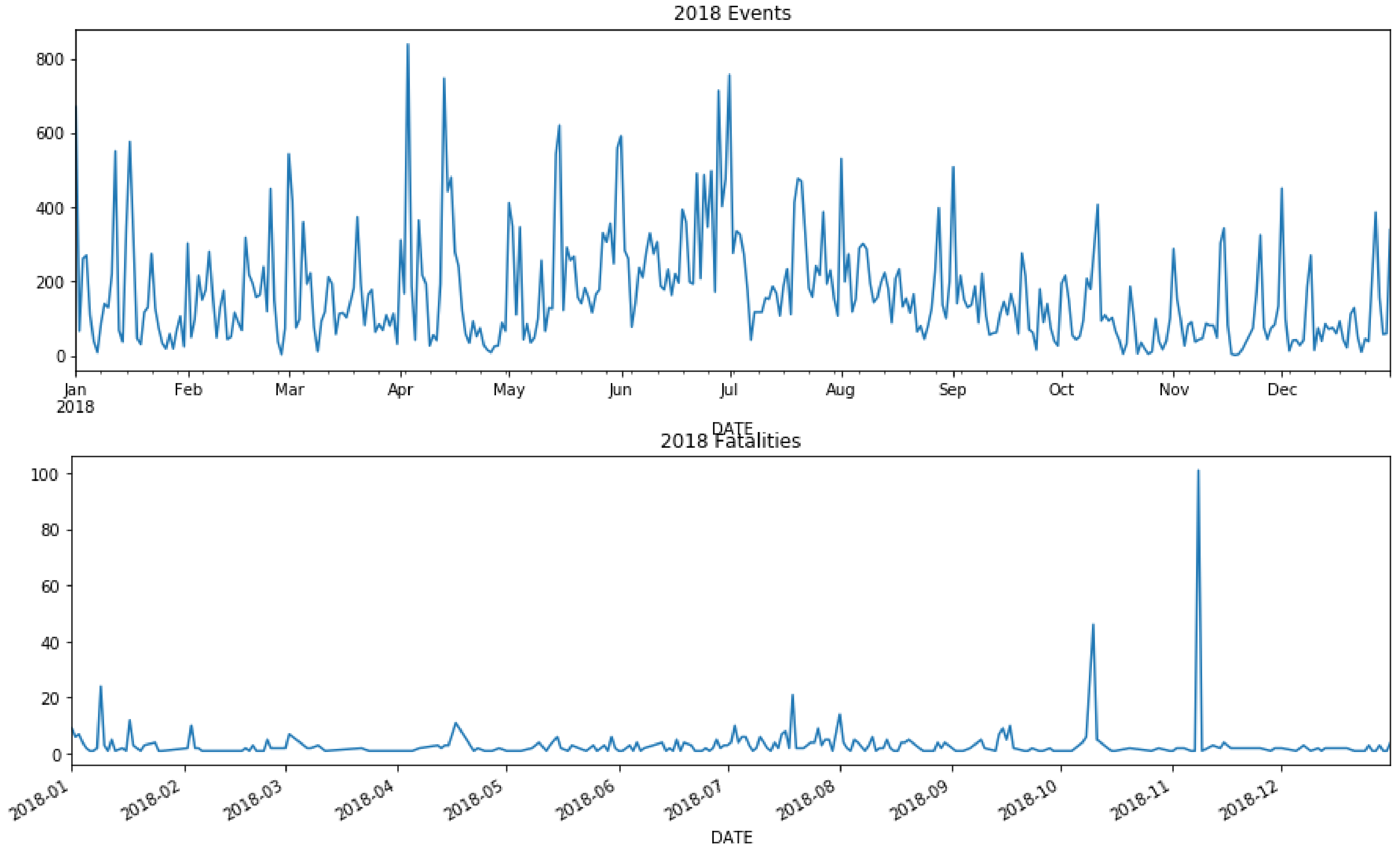
a. The occurrence of storms and other significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, and/or disruption to commerce;

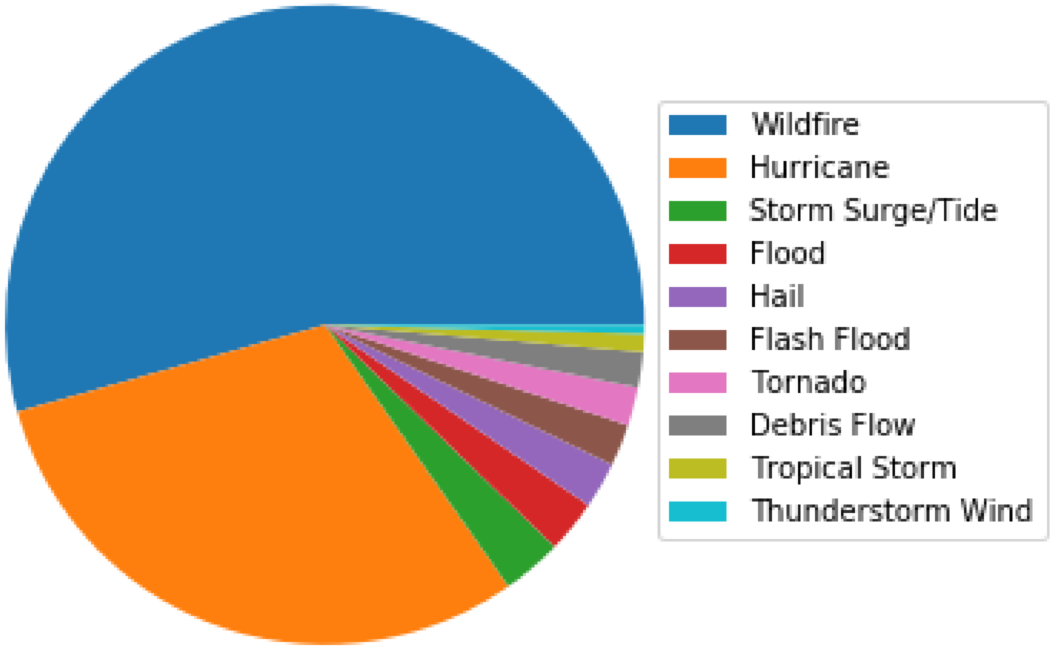
b. Rare, unusual, weather phenomena that generate media attention, such as snow flurries in South Florida or the San Diego coastal area; and

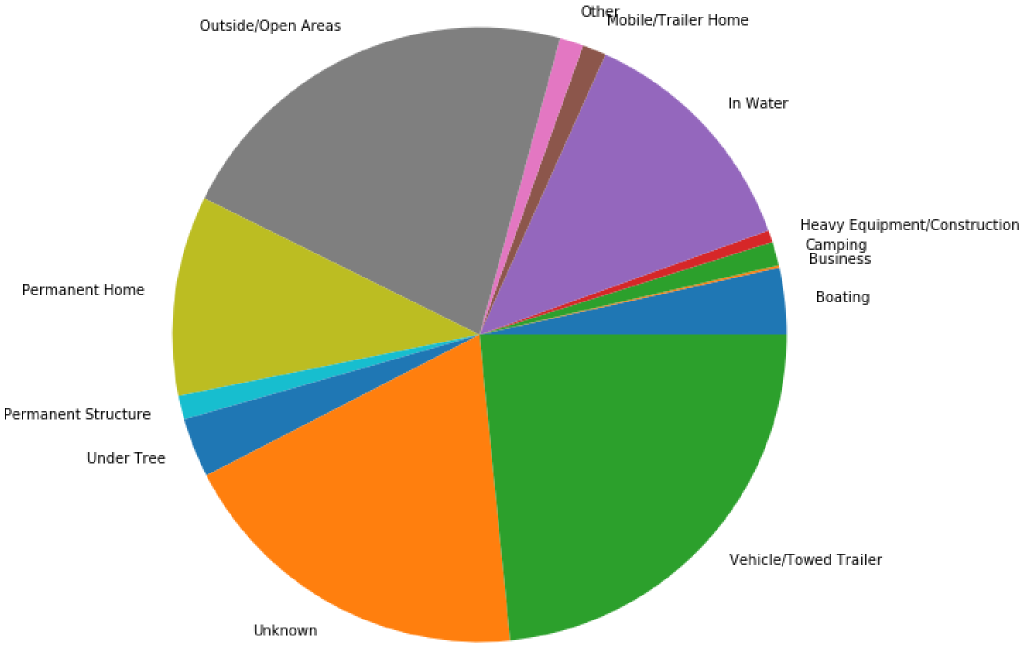
c. Other significant meteorological events, such as record maximum or minimum temperatures or precipitation that occur in connection with another event.

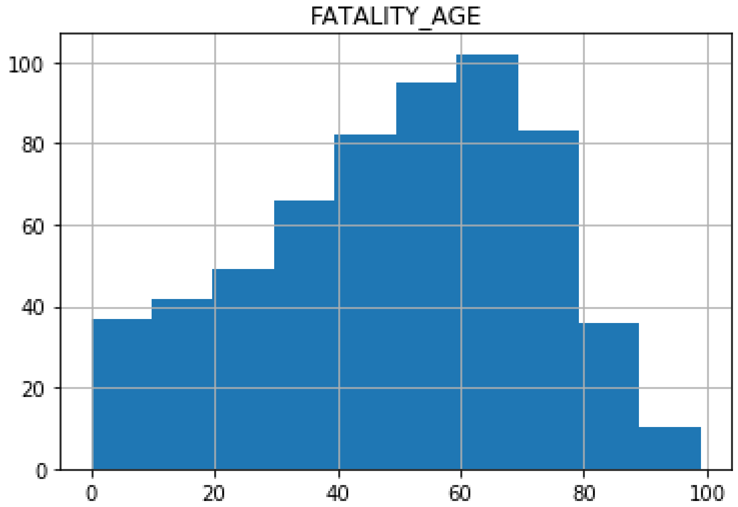
The database currently contains data from **January 1950 to July 2019**, as entered by NOAA's National Weather Service (NWS). Due to changes in the data collection and processing procedures over time, there are unique periods of record available depending on the event type. NCEI has performed data reformatting 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.

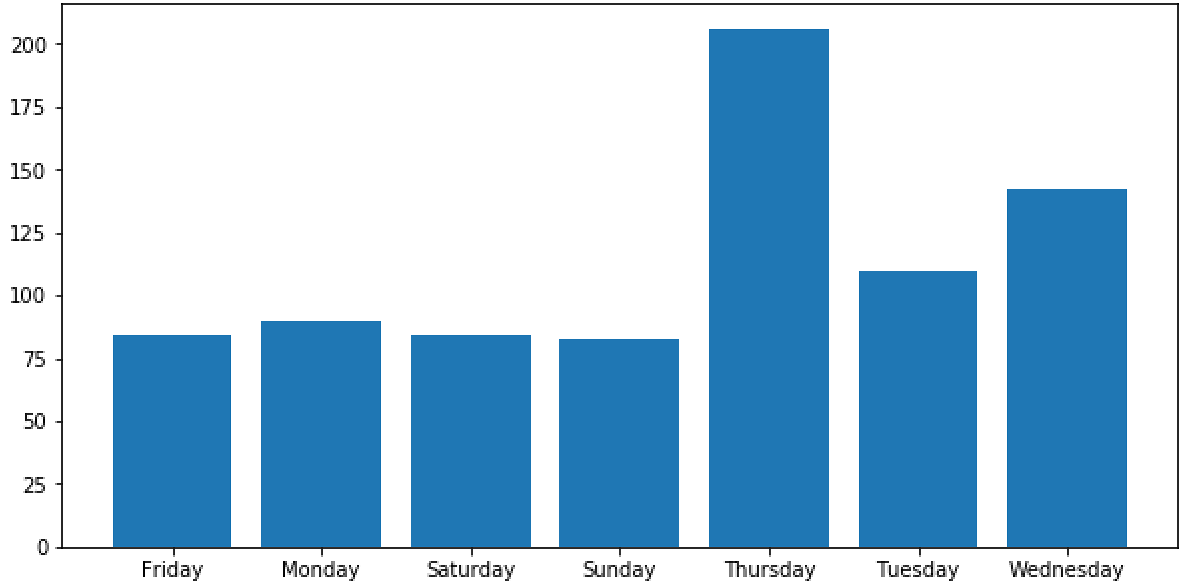
In 2018, 62351 Weather Events , 799 fatalities >$35 billion in property and crop losses




Wildfire
Hurricane
Storm Surge/Tide
Flood
Hail
Flash Flood
Tornado
Debris Flow
Tropical Storm
Thunderstorm Wind


Outside/Open Areas
Other
Mobile/Trailer Home
In Water
Heavy Equipment/Construction
Camping
Business
Boating
Permanent Home
Permanent Structure
Under Tree
Vehicle/Towed Trailer
Unknown


FATALITY_AGE
100
80
60
40
20
0
0
20
40
60
80
100


200
175
150
125
100
75
50
25
0
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday

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

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%

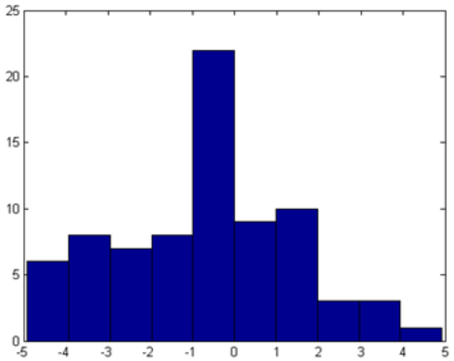
Distributions oriented: “ how close am I?”

- 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

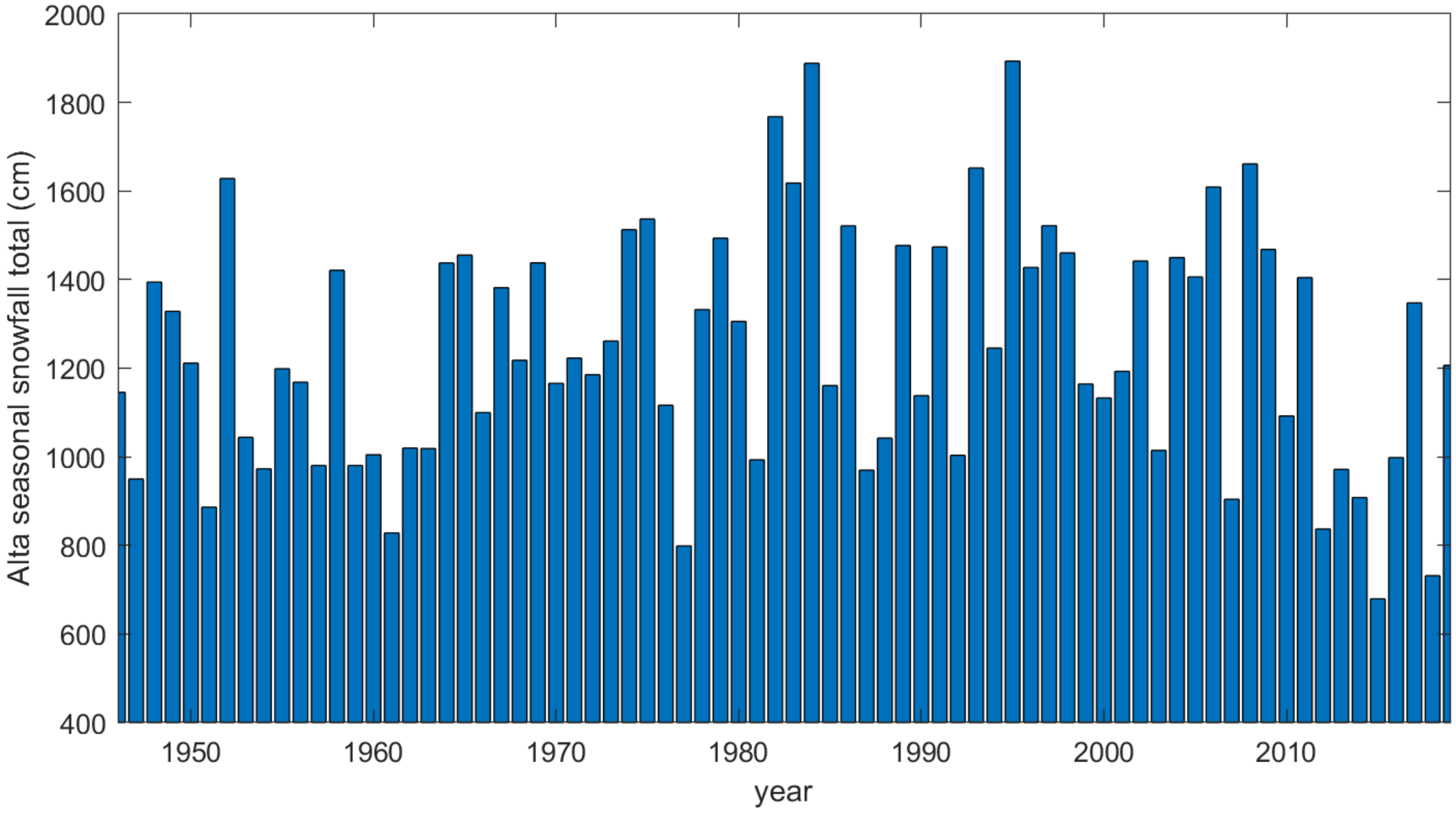


		Observed	Observed	Observed	Error Marginal totals
		≤3 m/s	3-6 m/s	≥6 m/s	
Error	≤ -2 m/s	0	10	11	21
Error	± 2 m/s	22	20	7	49
Error	> 2 m/s	0	7	0	7
	Observed Marginal totals	22	37	18	77

- 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 $Pr\{E_2 | E_1\}$:

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

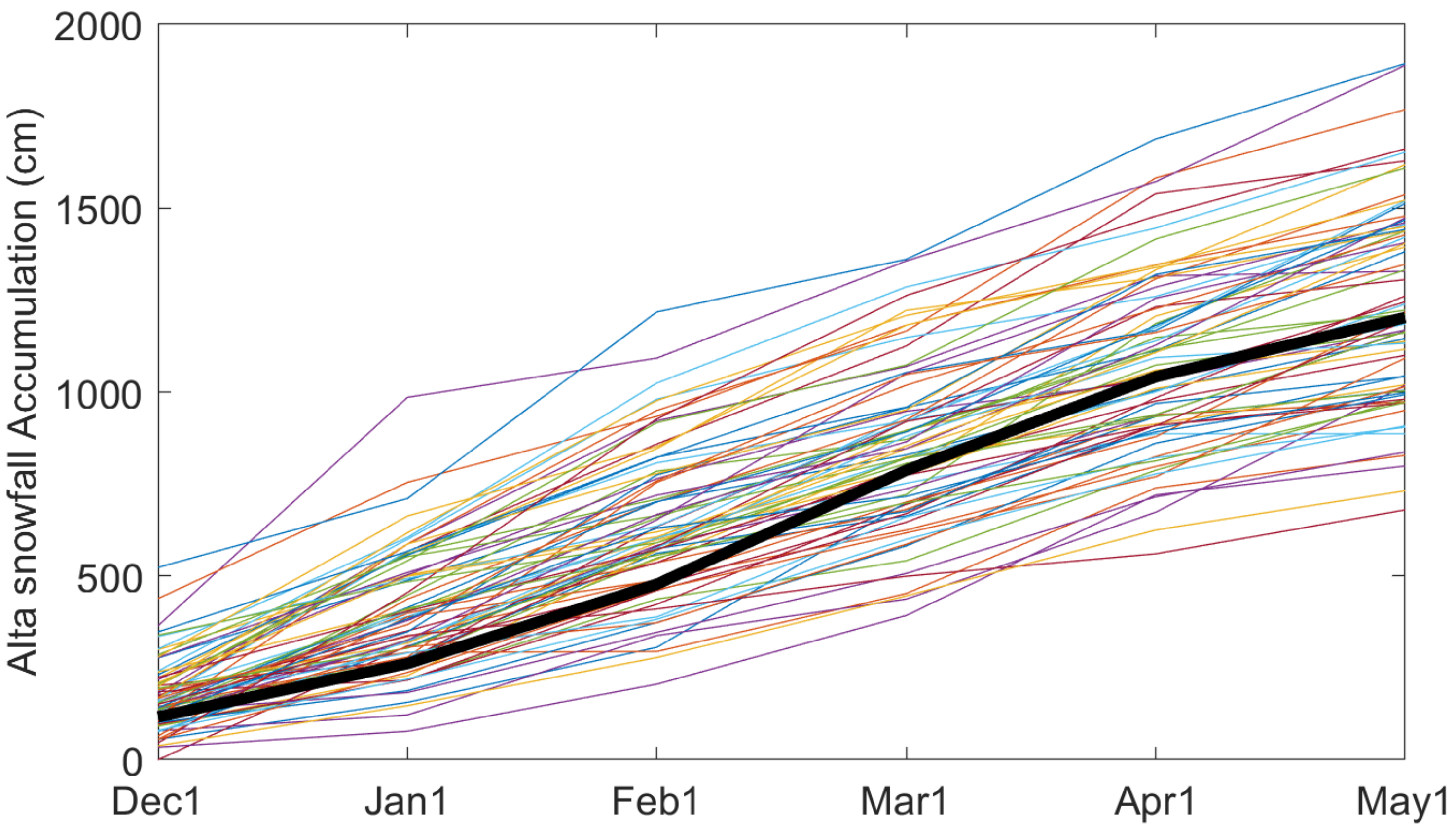
Alta Snowfall Seasonal Totals



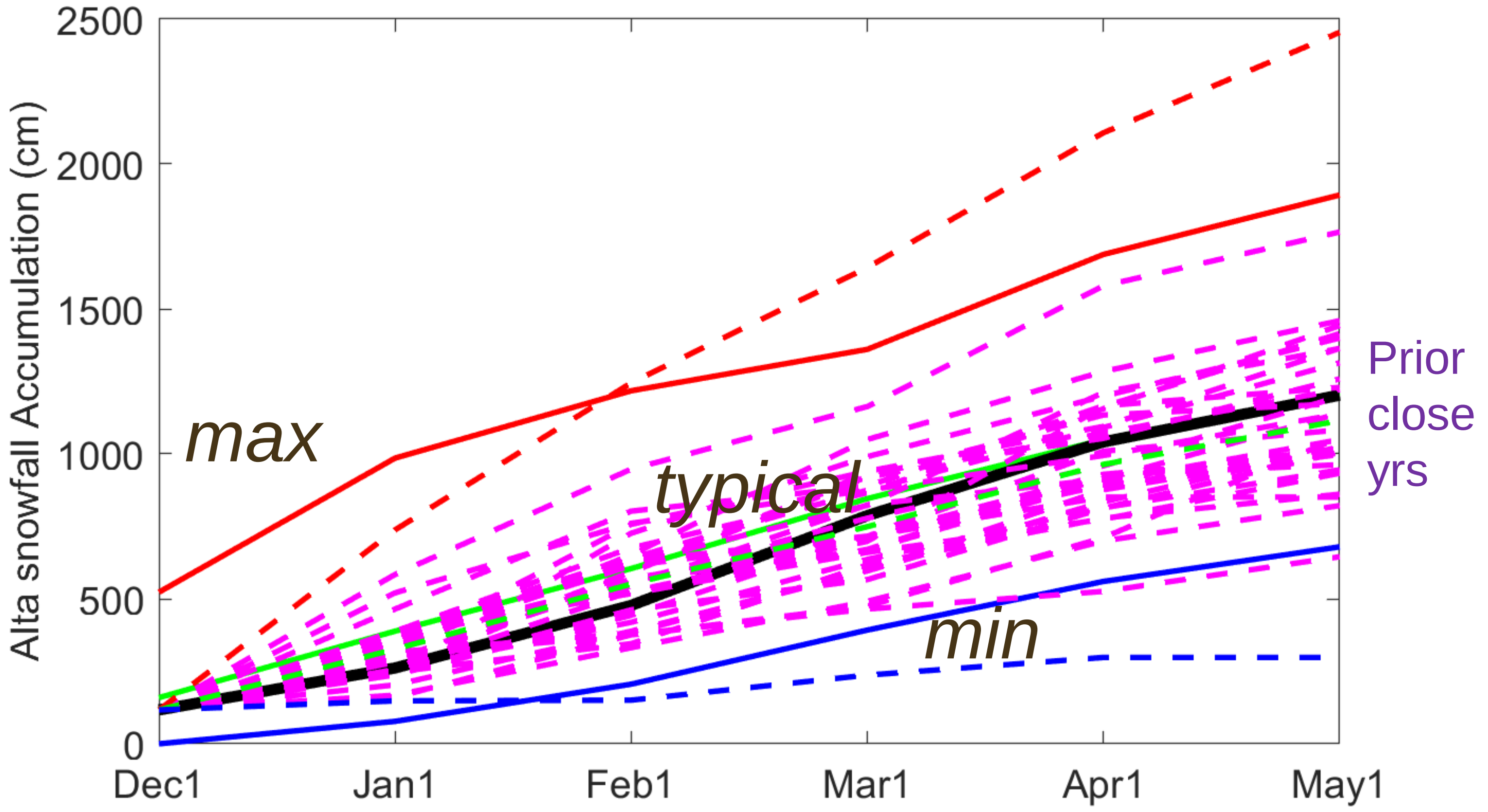
What can we say about estimating this winter's snow total will be?

- What physically is happening?
- Could we use last winter's snow total to predict this winter's?
  - Persistence from one year to next
- What about the amount of snow right now?
  - Persistence from one month to the next…

Atla Snowfall Accumulation Each Winter



How Much Snow Might Accumulate During this Season (19-20)at Alta?



Predict May 1 Snowfall from Dec 1 Snowfall

Case 1. Predictor: Dec1 total snowfall (cm)					
Predictand: May 1 Total snowfall at Alta (cm)		Below (E_1)	Near (E_2)	Above (E_3)	M Marginal Totals
	Below (M_1)	14	10	1	25
	Near (M_2)	7	7	10	24
	Above (M_3)	4	7	14	25
	E Marginal Totals	25	24	25	74

Predict May 1 Snowfall from Dec 1 Snowfall

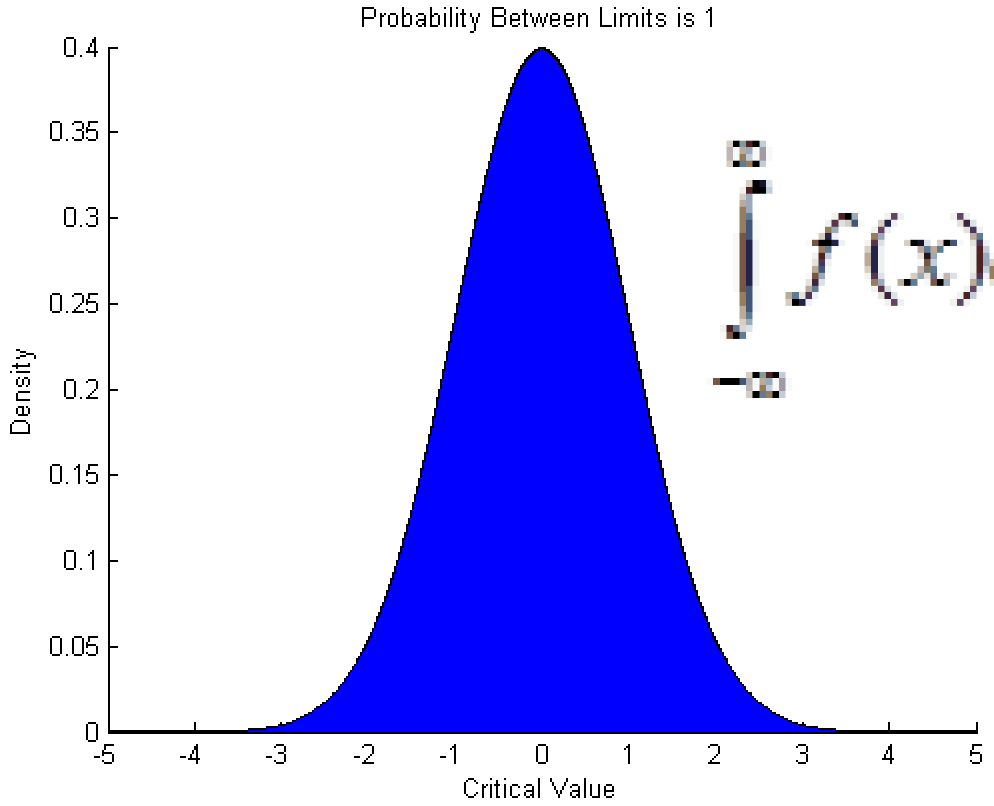
$Pr\{E_1\}$	**25/74=34%**	**$Pr\{M_3 \mid E_3\}$**	**14/25=56%**
$Pr\{E_1 \cap E_2\}$	0	$Pr\{E_1 \mid M_1\}$	14/25=56%
$Pr\{E_1 \cap M_1\}$	14/74=19%	$Pr\{E_2 \mid M_1\}$	10/25=40%
$Pr\{E_1 \cap M_3\}$	4/74=5%	$Pr\{E_3 \mid M_1\}$	1/25=4%
$Pr\{M_1 \mid E_1\}$	14/25=56%	$Pr\{E_3 \mid M_3\}$	14/25=56%
$Pr\{M_3 \mid E_1\}$	4/25=16%	$Pr\{E_1 \cap M_1\}$ IF random	100/9=11%
$Pr\{M_3 \mid E_2\}$	7/24=29%	% May 1 total same as Predictor	35/74=47%

Predict May 1 Snowfall from Last Year's May 1 Snowfall

$Pr\{E_1\}$		$Pr\{M_3 \mid E_3\}$	
$Pr\{E_1 \cap E_2\}$		$Pr\{E_1 \mid M_1\}$	
$Pr\{E_1 \cap M_1\}$		$Pr\{E_2 \mid M_1\}$	
$Pr\{E_1 \cap M_3\}$		$Pr\{E_3 \mid M_1\}$	
$Pr\{M_1 \mid E_1\}$		$Pr\{E_3 \mid M_3\}$	
$Pr\{M_3 \mid E_1\}$		$Pr\{E_1 \cap M_1\}$ IF random	
$Pr\{M_3 \mid E_2\}$		% May 1 total same as Predictor	

Empirical vs. Parametric Distributions

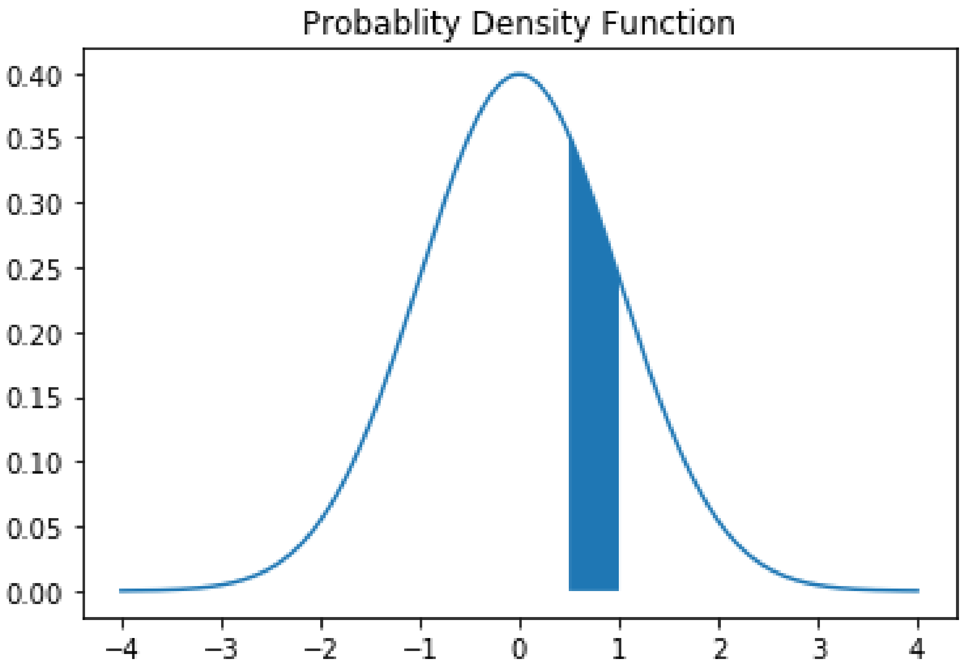
- Parameteric distributions:
  - Theoretical approach to define populations with known properties
  - Can be defined by a function with couple parameters and assumption that population composed of random events



$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Random Continuous Variable x

- f(x) probability density function (PDF) for a random continuous variable x
- f(x)dx incremental contribution to total probability

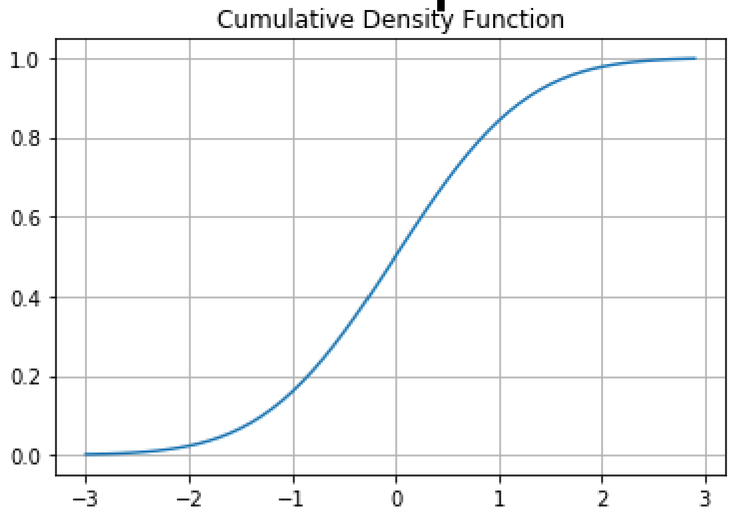


$$\int_{.5}^{1} f(x)dx$$

Cumulative Density Function of Continuous Variable

- F(X)- total probability below a threshhold
- F(0) = 50%
- F(.66) = 75%
- X(F) – quantile function- value of random variable corresponding to particular cumulative probability
- X(75%) = 0.66

$$F(X) = \Pr\{x \le X\} = \int_{-\infty}^{X} f(x)dx$$



Gaussian Parametric Distribution

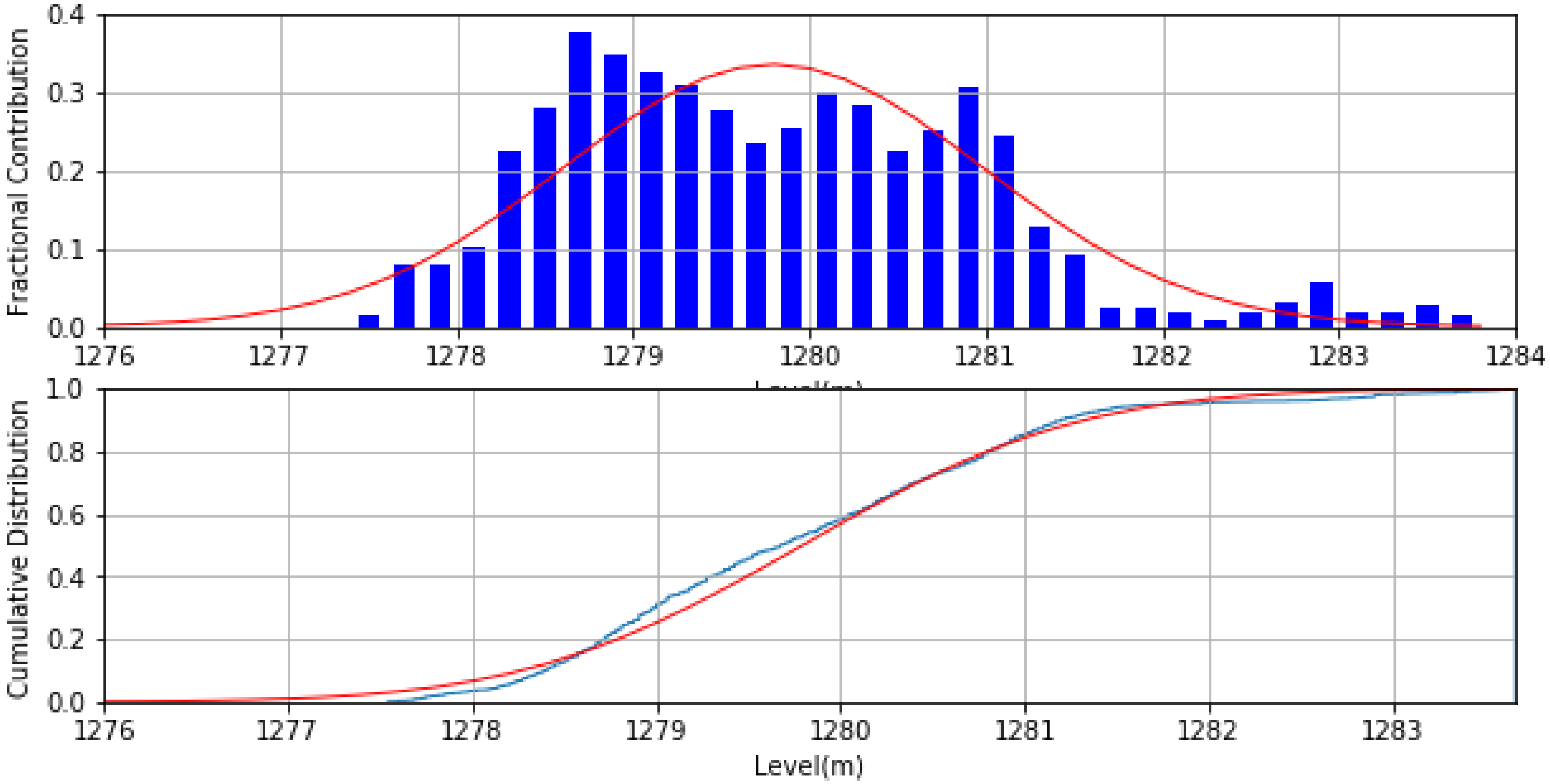
- PDF $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$
- CDF $$F(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{X} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

- Two parameters define Gaussian distribution: μ and σ
- Nothing magic or "normal" about the Gaussian distribution- it is a mathematical construct
- Previous examples are for a Gaussian distribution only

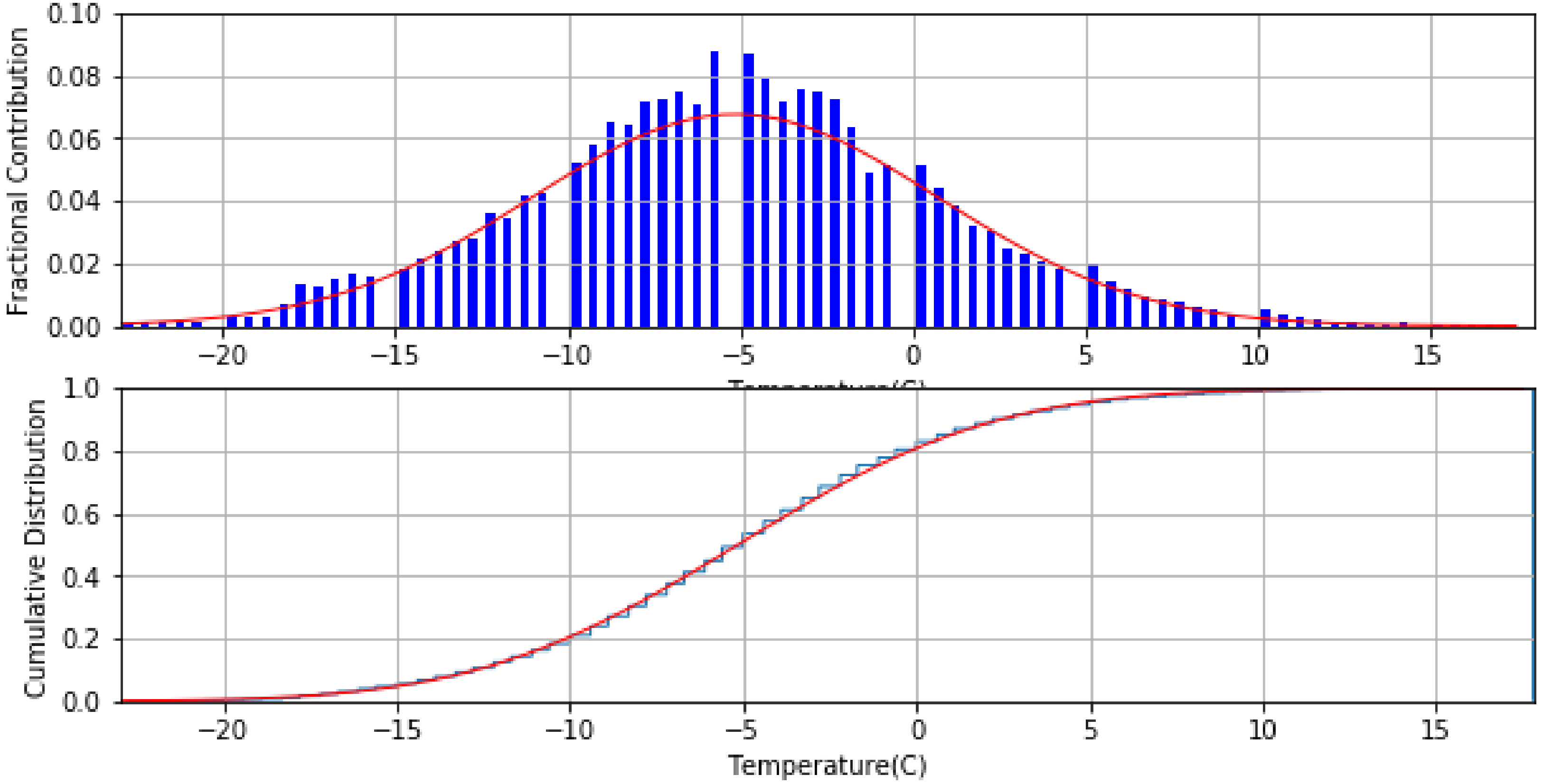
Using parametric distributions

- Generate an empirical cumulative probability (CDF)
- See if there is a good match between the empirical CDF and a particular parametric distribution
- Use the parameters from that parametric distribution to estimate the probabilities of values above below a threshhold or extreme events

Great Salt Lake Level
Gaussian fit using sample mean & sample estimate of population std dev



Alta Winter Temperatures with Gaussian fit using mean, σ

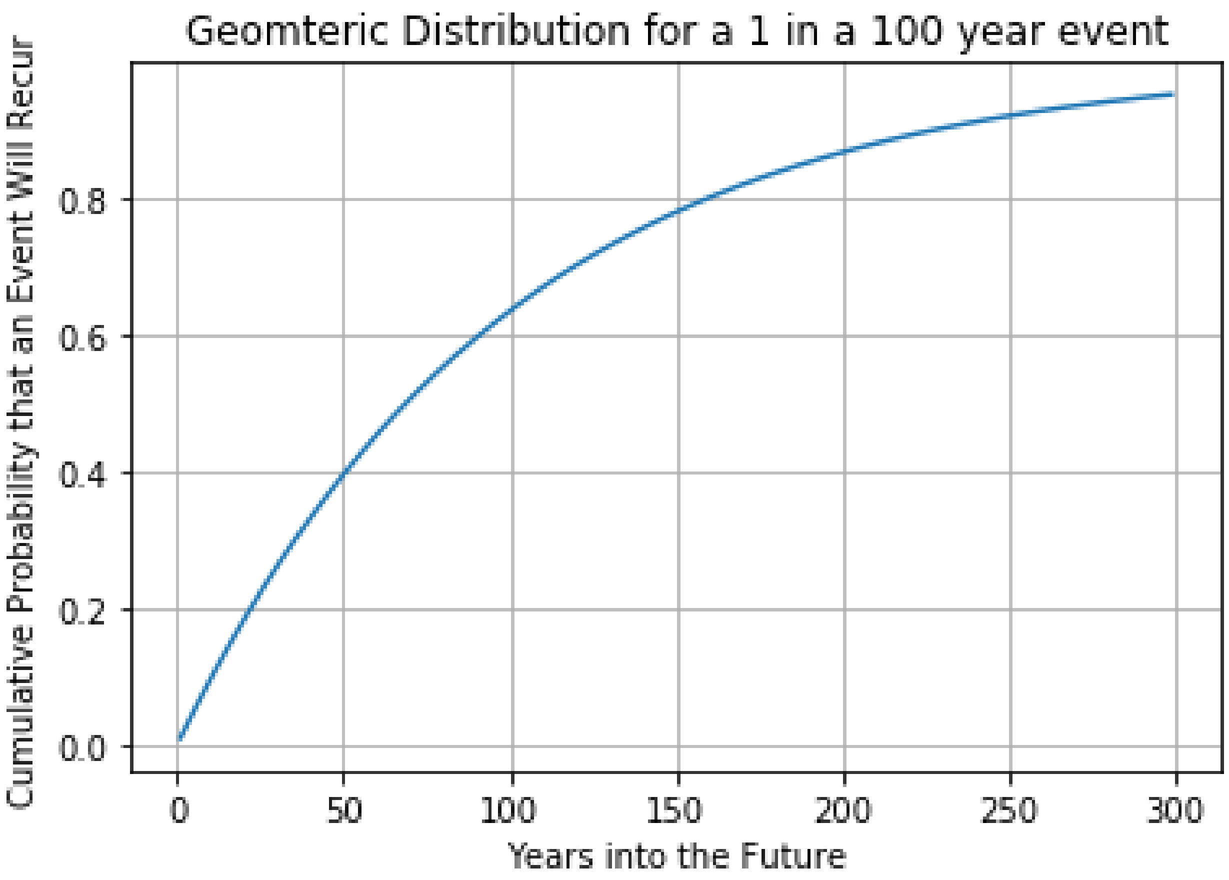


Using normspec: normal density plot

- 68.3% between -1 and 1
- 95.% between -2 and 2
- 2.3% of time variable explained by Gaussian distribution > 2 std dev of mean

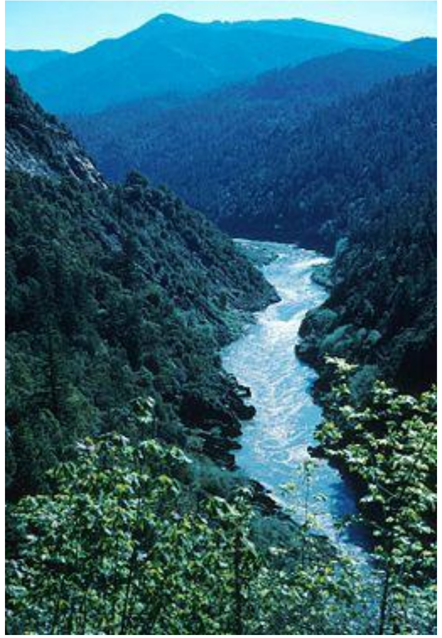
Geometric Distribution

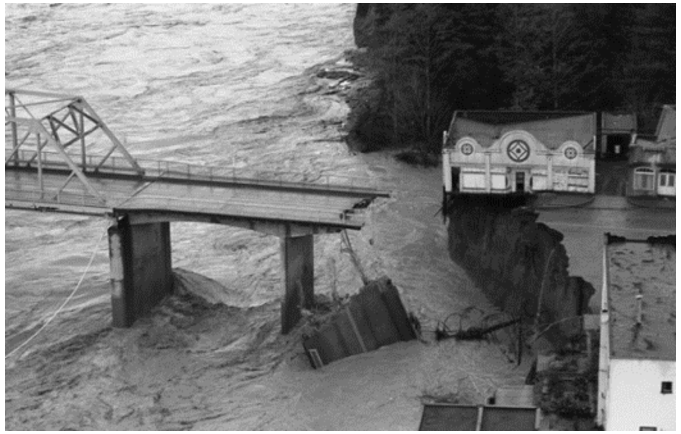
- Estimating how likely rare events can happen by chance
- Pr{0.01}- probability of a 1 in 100 year event
- probability for the next event to happen in 1, 10, 30, 100, 200, 300 years
- 63% chance in next 100 years
- 12% chance not until 200 years
- This is nothing “real”, just one of many assumptions you might choose



Klamath River, northern CA

http://water.weather.gov/ahps2/hydrograph.php?wfo=eka&gage=klmc1





Flood Categories (in feet)

Major Flood Stage:	46
Moderate Flood Stage:	42
Flood Stage:	38
Action Stage:	30

Historic Crests

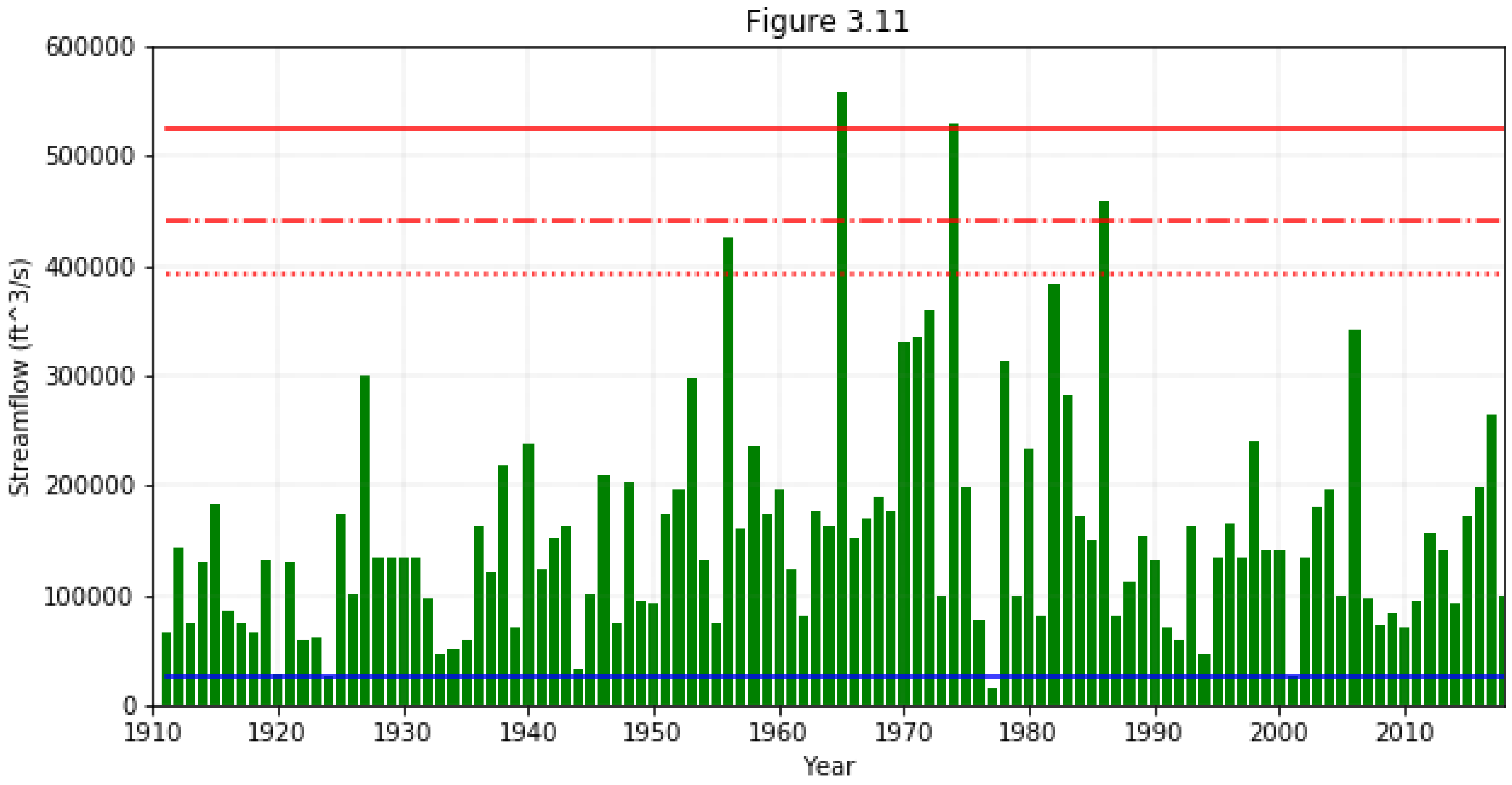
(1) 61.29 ft on 12/23/1964
(2) 47.12 ft on 12/31/2005
(3) 43.80 ft on 01/01/1997
(4) 41.64 ft on 02/10/2017 **(P)**
(5) 40.52 ft on 12/29/2005
Show More Historic Crests

(P): Preliminary values subject to further review.

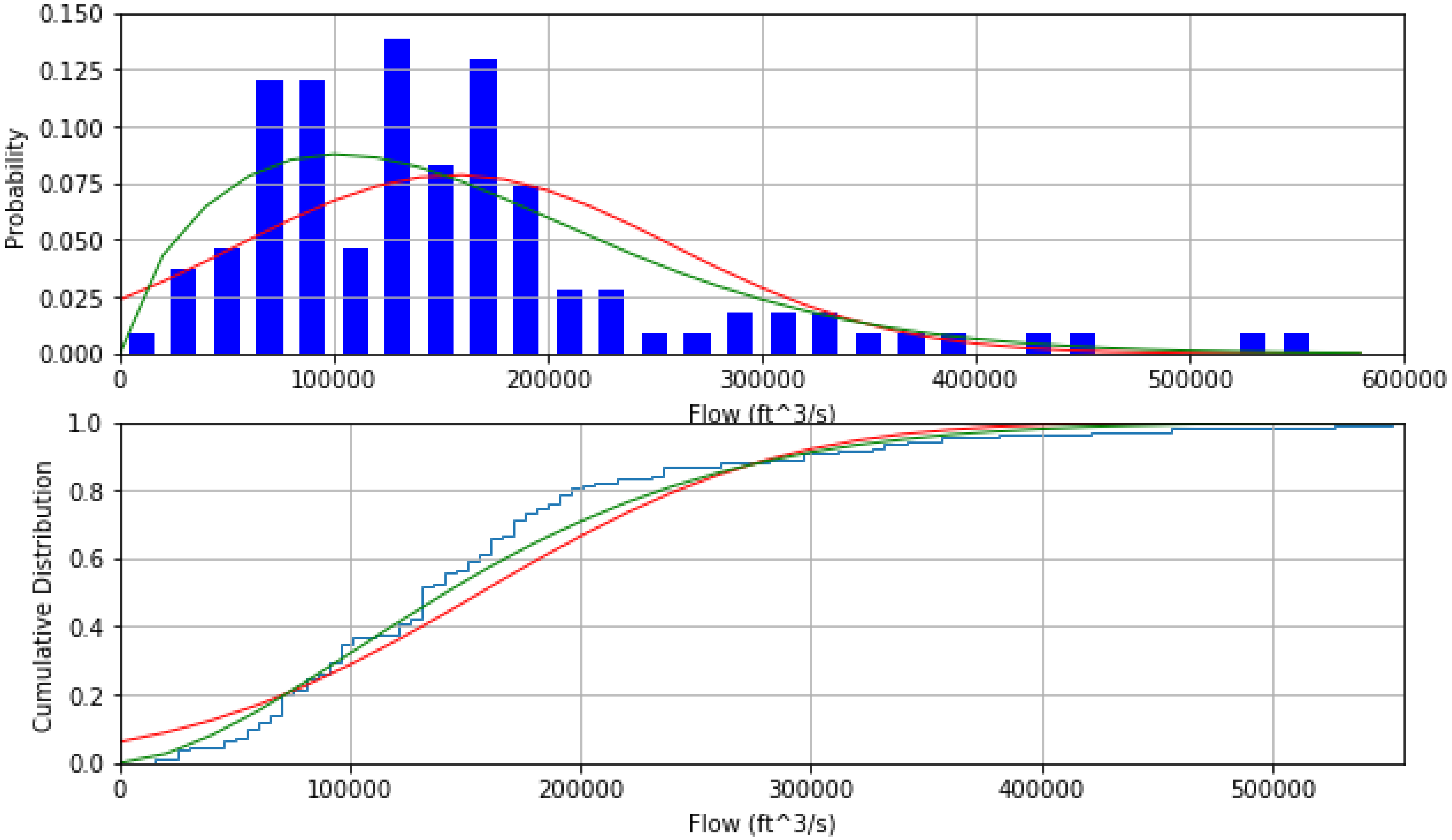
Recent Crests

(1) 41.64 ft on 02/10/2017 **(P)**
(2) 25.51 ft on 03/10/2014
(3) 30.83 ft on 12/02/2012
(4) 32.33 ft on 03/31/2012
(5) 25.82 ft on 12/29/2010

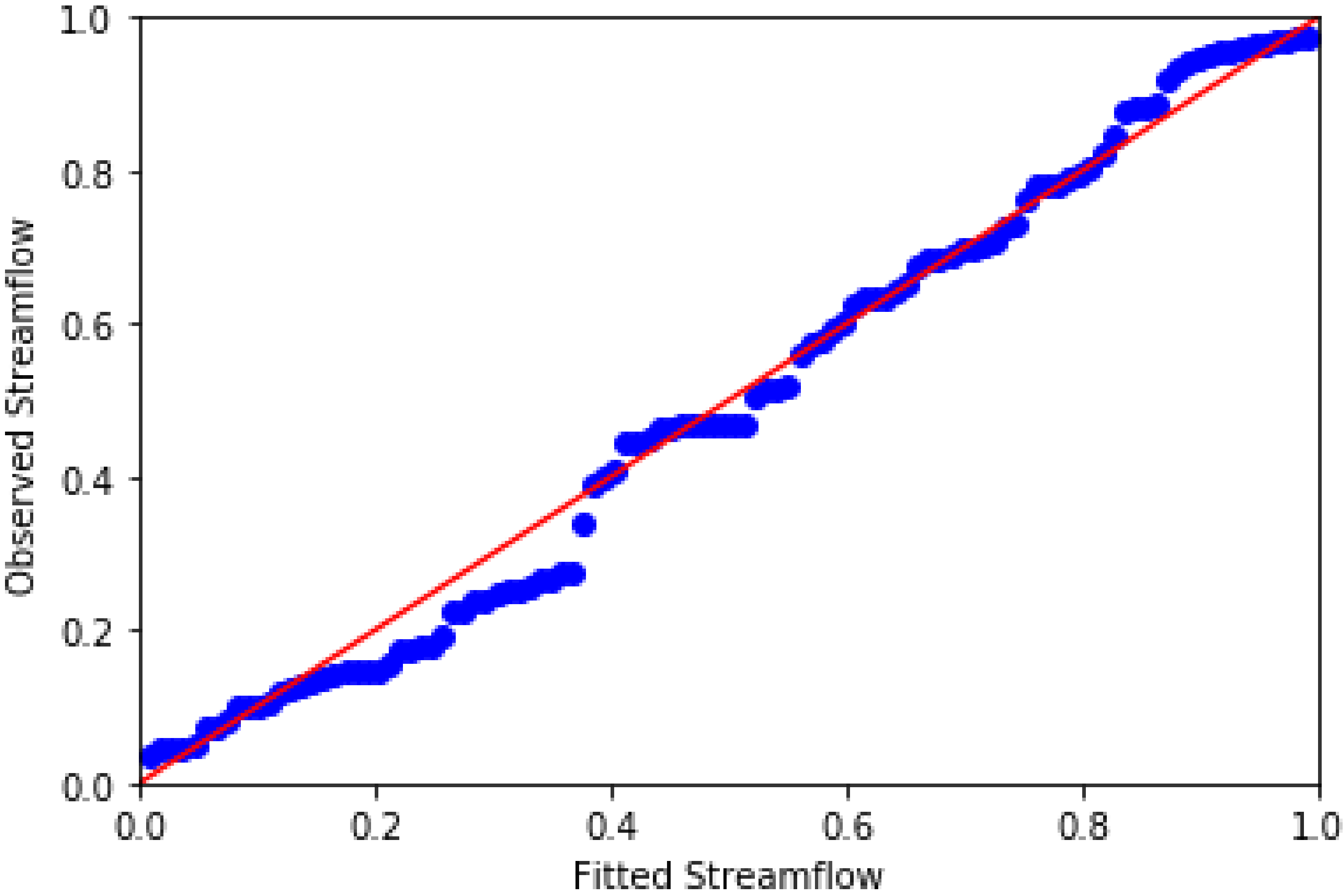
Klamath River Streamflow



Klamath River Streamflow: Gaussian & Weibull parametric fits



Klamath streamflow

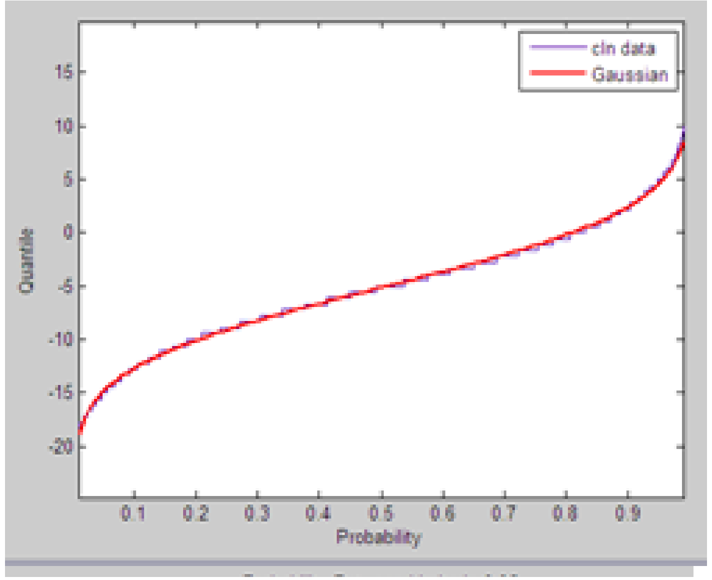


What you should be doing

- Read all of Chapter 3 Notes
- Storm Events Assignment 6 Due Nov. 20
- Turn in today’s in class assignment Nov. 21

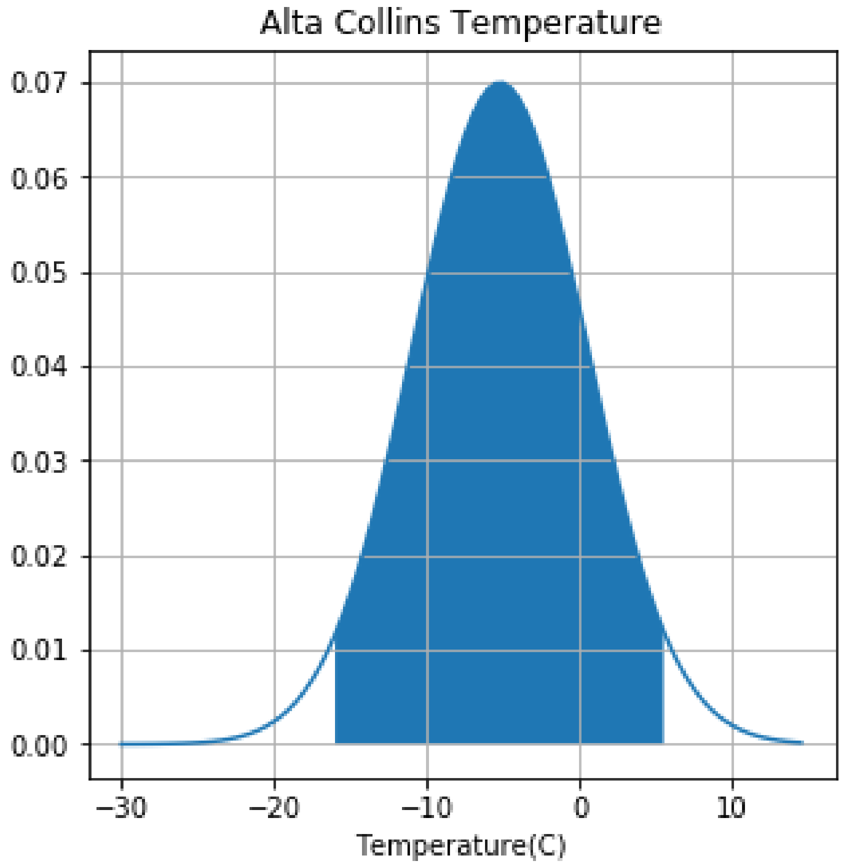
Hypothesis Testing

- Alta temperature:
- Empirically: probability of temperature less than -15C is low
- Empirical estimates:
  - Mean= -5.1C
  - Std dev = 5.9C
- What are chances of getting temp of -20C IF this was a population of random numbers with that mean and std dev?



Null hypothesis

- Null hypothesis: Temp of -20C does not differ significantly from mean of -5.1C
- 95% of time, random value would be within -16 and 6C
- So 5% of time, random value would be outside this range
- REJECT the null hypothesis accepting a 5% risk that we are rejecting the null hypothesis incorrectly
- If null hypothesis: Temp of -15C does not differ significantly from mean of -5.1C
- CANNOT reject the null hypothesis since 95% of the time the value could be within -16 and 6C



Warning

Don't use language such as:

- the results are significant at the 5% level

You can only state:

- Null hypothesis is rejected as too unlikely to be true with a risk of 5%

Collins: Confidence Intervals