University of California, Davis
Department of Land, Air and Water Resources

Expected Monetary Value

ESM-121 Water Science and Management

Samuel Sandoval Solis, PhD
Assistant Professor

BASIC CONCEPTS

- The mean

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

	Precip		Area(mi²)	
A1	11	×	50	+
A2	15	×	25	+
A3	20	×	10	
	46		85	

- Weighted mean

$= X \; P(x)$

$\bar{P} = \frac{46}{3} = 15.3''$

$$\boxed{\bar{X} = \sum_{i=1}^{n} \bar{X}_i \frac{n_i}{n}}$$

$\bar{P} = \frac{11*50+15*25+20*10}{85}$

$\bar{P} = 13.2''$

- Median

median $(P_{0.50}) = X_{(n+1)/2}$ when n is odd, and

median $(P_{0.50}) = \frac{1}{2}(X_{(n/2)} + X_{(n/2)+1})$ when n is even.

MEAN VS MEDIAN

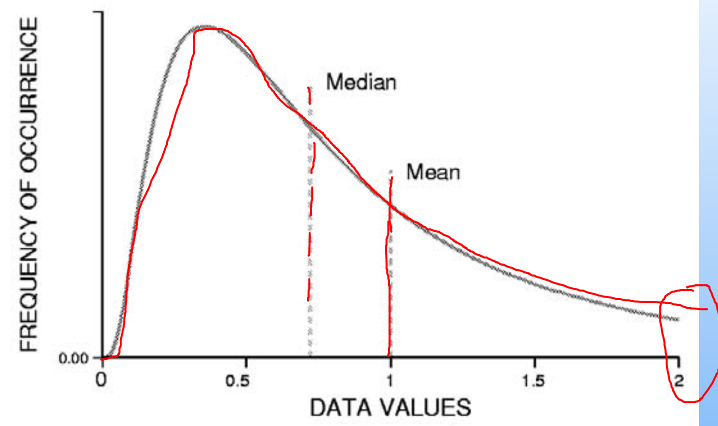


Figure 1.1 Density Function for a Lognormal Distribution

MEAN VS MEDIAN

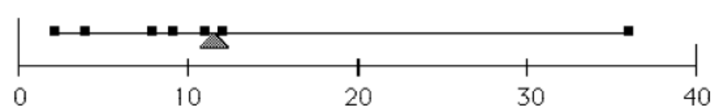


Figure 1.3a The mean (triangle) as balance point of a data set.

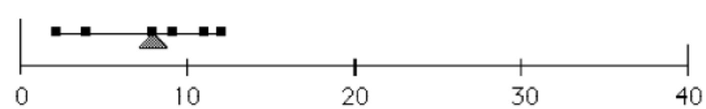


Figure 1.3b Shift of the mean downward after removal of outlier.

Example 1:

			Med
(a)	2 4 8 9 11 11 12	$\bar{X}$ = 8.1	$P_{.50}$ = 9
(b)	2 4 8 9 11 11 120	$\bar{X}$ = 23.6	$P_{.50}$ = 9

MEAN VS MEDIAN

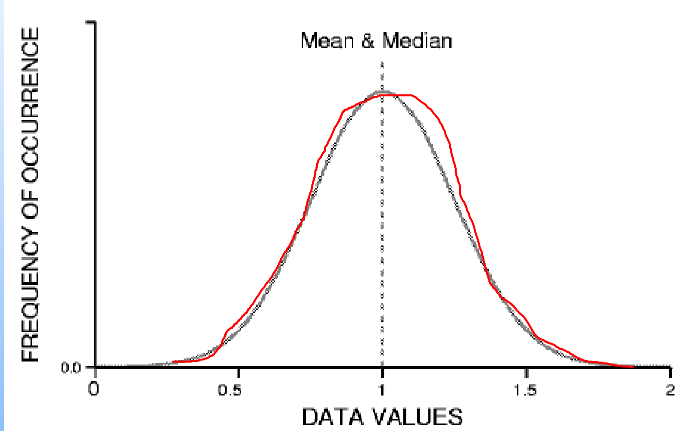


Figure 1.2 Density Function for a Normal Distribution

BASIC CONCEPTS

- Variance and Standard Deviation

$$s^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{(n-1)} \qquad s = \sqrt{s^2}$$

- Interquartile Range

IQR = $P_{0.75} - P_{0.50}$

- Median Absolute Deviation

MAD (X_i) = median $|d_i|$, where $d_i = X_i -$ median (X_i)

STD. DEVIATION VS IQR VS MAD

data	2	4	8	9	11	11	12	IQR = 11 – 4 = 7
$(X_i - \overline{X})^2$	37.2	16.8	0.01	0.81	8.41	8.41	15.2	$s^2 = (3.8)^2$
$\lvert d_i = X_i - P_{.50} \rvert$	7	5	1	0	2	2	3	MAD=median $\lvert d_i \rvert$=2
data	2	4	8	9	11	11	120	IQR = 11 – 4 = 7
$(X_i - \overline{X})^2$	37.2	16.8	0.01	0.81	8.41	8.41	12,522	$s^2 = (42.7)^2$
$\lvert d_i = X_i - P_{.50} \rvert$	7	5	1	0	2	2	111	MAD=median $\lvert d_i \rvert$=2

HISTOGRAMS

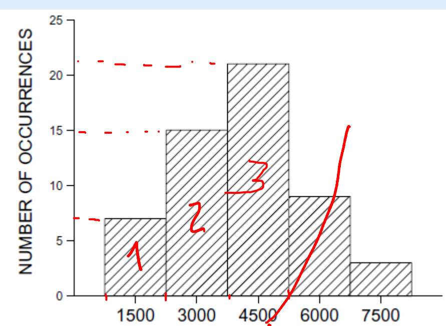
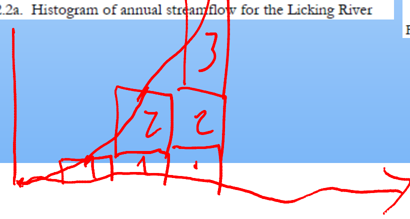


Figure 2.2a. Histogram of annual streamflow for the Licking River



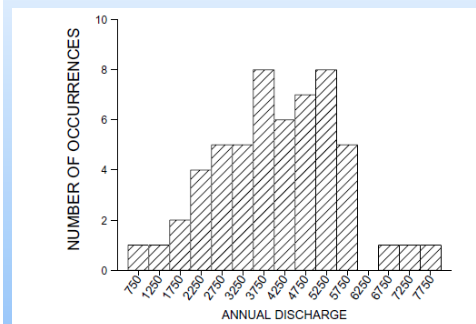


Figure 2.2b. Second histogram of same data, but with different interval divisions.

QUARTILE PLOTS

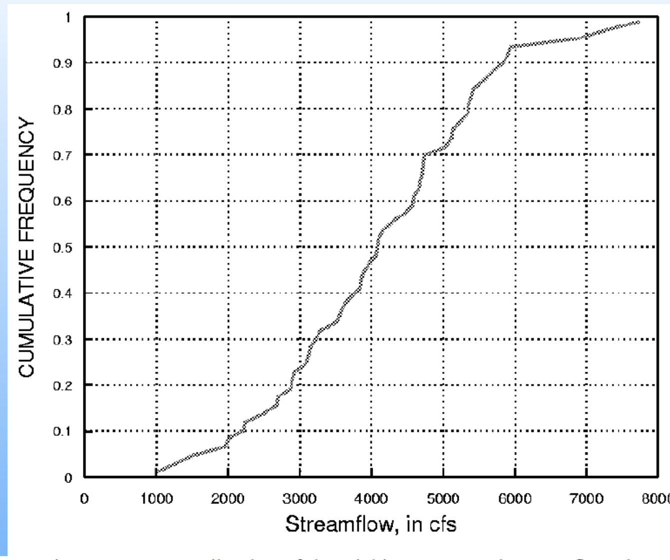
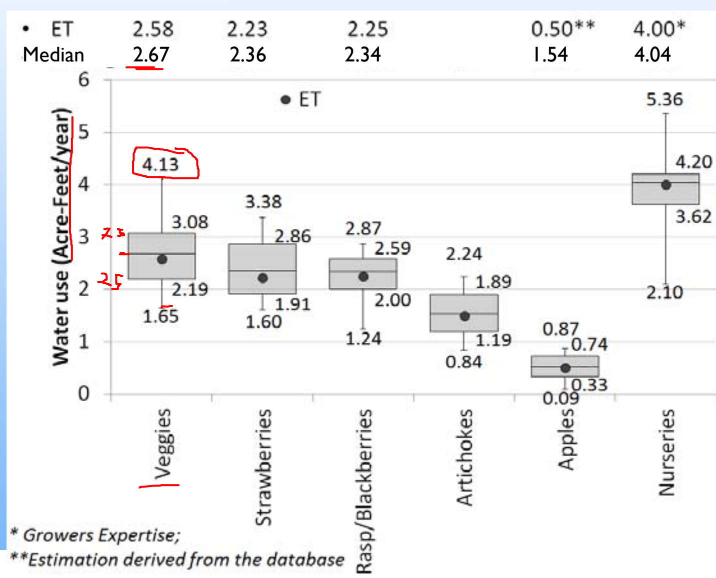


Figure 2.4 Quantile plot of the Licking R. annual streamflow data

BOXPLOTS

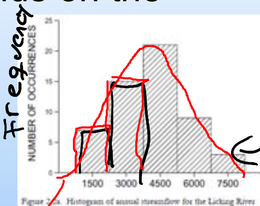


RANDOM VARIABLES (RV) PROBABILITIES VARIABLES

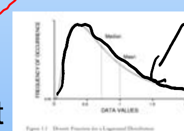
Data Freq

- Function (X) whose value (x) depends on the outcome of a chance event

Frequency



- Discrete RV (Data
  - Takes on values from a discrete set
    - # of years until a certain flood stage returns
    - # of times reservoir storage drops below a level



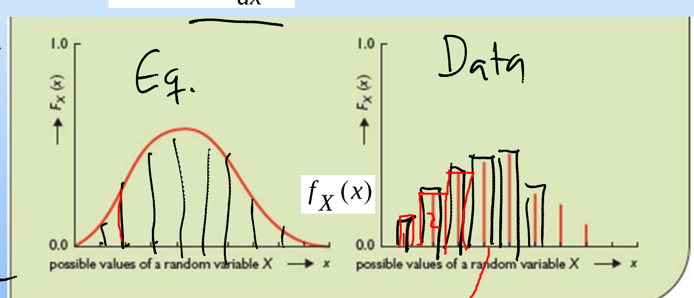
Formula Eq

- Continuous RV (Eqn)
  - Takes on values from a continuous set
    - e.g., Rainfall, Streamflow, Temperature, Concentration

PROBABILITY DENSITY (MASS) FUNCTIONS (PDF)

Probability density function	Probability mass function
$f_X(x) = \dfrac{dF_X(x)}{dx}$	$p_X(x_i) = \Pr[X = x_i]$

Relat. Freq 0-1



Continuous RV

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

$$F_X(x) = \int_{-\infty}^{x} f_X(s)ds$$

Discrete RV

CUMULATIVE DISTRIBUTION FUNCTIONS (CDF)

Add PDF

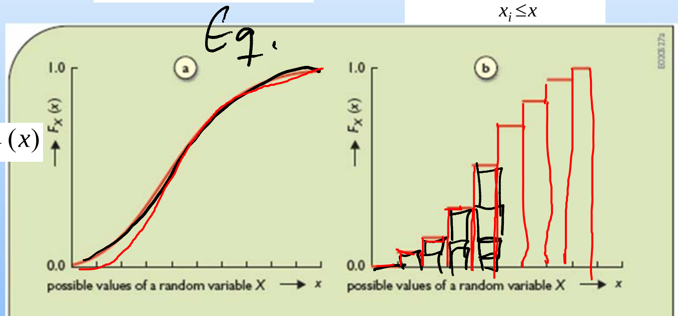
Cummulative

$F_X(x) = \Pr\{X \le X\}$

$F_X(x) = \sum_{x_i \le x} p_X(x_i)$

Eq.

$F_X(x)$



Continuous RV

Discrete RV

EXPECTATION

EV = $V_1 * P_1 + V_2 * P_2 + \ldots + V_n P_n$

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)dx \qquad X \text{ is continuous}$$

$$E[g(X)] = \sum_{x_i} g(x_i) p_X(x_i) \qquad X \text{ is discrete}$$

Note (expected value of X – Mean of X)

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x)dx$$

$, L, CY,

$$\boxed{E[X] = \sum_i x_i p_X(x_i)}$$

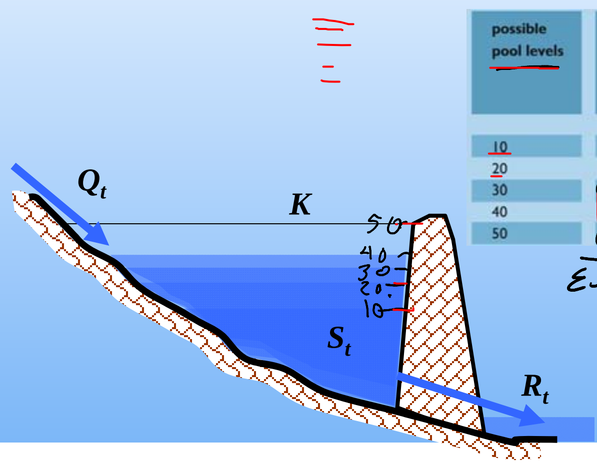
Probability

PRINCIPLE

Replacement of uncertain quantities by either expected, median or worst-case values can grossly affect the evaluation of project performance when important parameters are highly variable.

EXAMPLE

Elevation of reservoir water surface varies from year to year depending on the inflow and demand for water.



possible pool levels	probability of each level	recreation potential in visitor-days per day for reservoir with different pool levels
10	0.10	25
20	0.25	75
30	0.30	100
40	0.25	80
50	0.10	70

$\Sigma = 1.00$

EXAMPLE

1) → 2) PR

possible pool levels	probability of each level (3)	recreation potential in visitor-days per day for reservoir with different pool levels (2)
10	0.10	25
20	0.25	75
30	0.30	100
40	0.25	80
50	0.10	70

Average pool level

$$\bar{L} = E[L] = \sum_i l_i p_L(l_i)$$

$L_1(P_1) + L_2(P_2) + \dots + L_5(P_5) =$

$$\bar{L} = 10(0.10) + 20(0.25) + 30(0.30) + 40(0.25) + 50(0.10) = 30$$

EXAMPLE

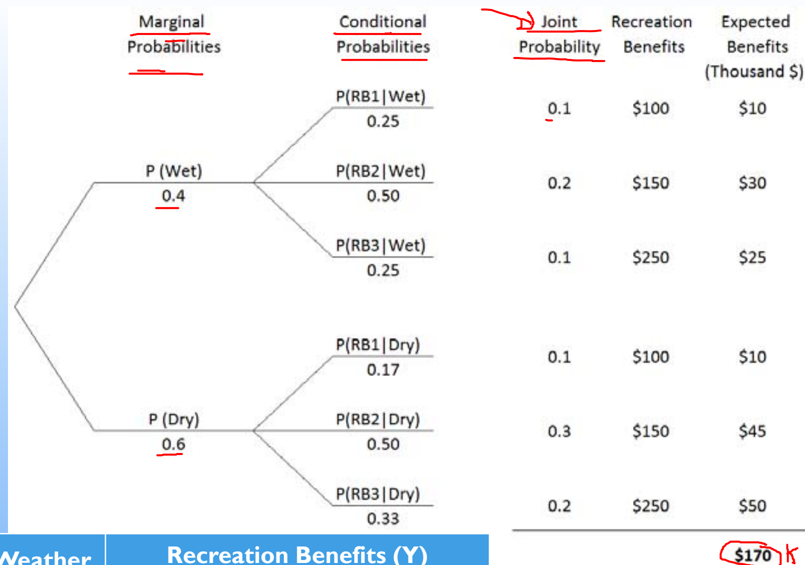
For example, the following matrix displays the probabilities of different weather conditions and of different recreation benefits levels obtained from use of a reservoir in a state park

$$\boxed{E[X] = \sum_i x_i p_X(x_i)}$$

Expected Benefits

Weather (X)	Recreation Benefits (Y)		
	RB_1	RB_2	RB_3
Wet	0.1	0.2	0.1
Dry	0.1	0.3	0.2

RB_1 = \$100K
RB_2 = \$150K
RB_3 = \$250K



Weather (X)	Recreation Benefits (Y)		
	RB_1	RB_2	RB_3
Wet	0.1	0.2	0.1
Dry	0.1	0.3	0.2

$$E[X] = \sum_i x_i p_X(x_i)$$

MULTIPLE RANDOM VARIABLES (RV)

The **joint distribution** of two RVs, X and Y:

$$F_{X,Y}(x,y) = \Pr\{X \le x \text{ and } Y \le y\}$$

And

EXAMPLE

For example, the following matrix displays the probabilities of different weather conditions and of different recreation benefits levels obtained from use of a reservoir in a state park

Relative Freq.

RB_1= $100K
RB_2= $150K
RB_3= $250K

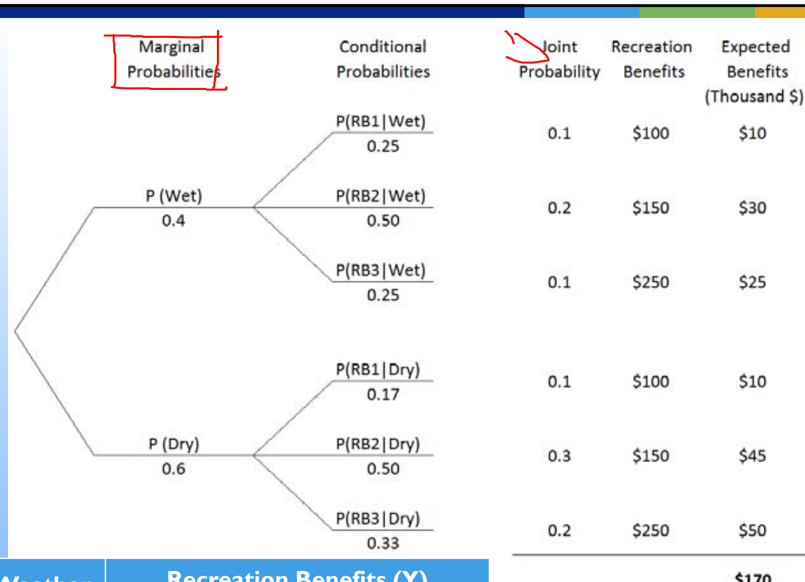
Weather (X)	Recreation Benefits (Y)		
	RB_1 Y_1	RB_2	RB_3
Wet X_1	0.1	0.2	0.1
Dry X_2	0.1	0.3	0.2

$\Sigma = 1$

INDEPENDENT RV

- If the distribution of RV X is not influenced by the value taken by RV Y, and vice versa, the RVs are *independent*
- For two independent RVs, the joint probability is the product of the separate probabilities.

$$\Pr\{X \le x \text{ and } Y \le y\} = \Pr\{X \le x\}\Pr\{Y \le y\}$$



Weather (X)	Recreation Benefits (Y)		
	RB_1	RB_2	RB_3
Wet	0.1	0.2	0.1
Dry	0.1	0.3	0.2

$$E[X] = \sum_i x_i p_X(x_i)$$

MARGINAL DISTRIBUTIONS

Two RVs X and Y can have a joint distribution

$F_{XY}(x,y)$

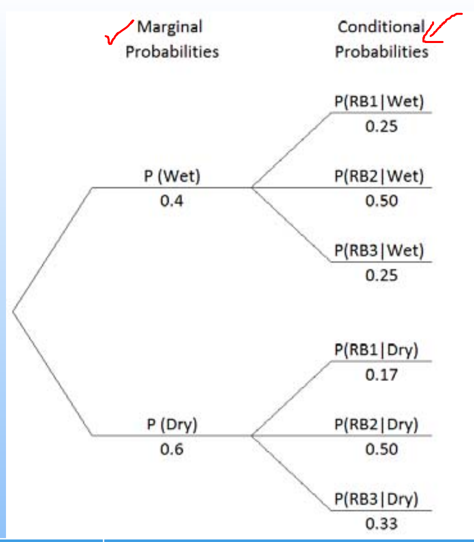
The **marginal distribution** of X (Y) is the distribution of X ignoring ~~Y~~

$$\lim_{y\to\infty} F_{XY}(x,y) = \Pr\{X \le x\} = F_X(x)$$

X ↓ Y1 Y2 Y3

Weather (X)	Recreation Benefits (Y)			
	RB_1	RB_2	RB_3	
Wet	0.1 +	0.2 +	0.1 =	Σ=0.4
Dry	0.1 +	0.3 +	0.2 =	Σ=0.6
Y	0.2 +	0.5 +	0.3 =	1.0

X	
wet	0.4
Dry	0.6
	Σ=1.0



Joint Probability	Recreation Benefits	Expected Benefits (Thousand $)
0.1	$100	$10
0.2	$150	$30
0.1	$250	$25
0.1	$100	$10
0.3	$150	$45
0.2	$250	$50
		$170

Weather (X)	Recreation Benefits (Y)		
	RB_1	RB_2	RB_3
Wet	0.1	0.2	0.1
Dry	0.1	0.3	0.2

$$E[X] = \sum_i x_i p_X(x_i)$$

CONDITIONAL DISTRIBUTIONS

Conditional distribution of X *given* that Y has taken on a particular value

$$F_{X|Y}(x|y) = \Pr\{X \le x \text{ given } Y = y\}$$

$$p_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)} = \frac{}{P(\text{wet})} = \frac{}{0.4}$$

Weather (X)	Recreation Benefits (Y)		
	RB_1	RB_2	RB_3
Wet	0.1	0.2	0.1
Dry	0.1	0.3	0.2

=0.4 (Wet), =0.6 (Dry); 0.1/0.6, 0.3/0.6, 0.2/0.6

wet: 0.25 RB_1, 0.50 RB_2, 0.25 RB_3

$P(RB_1 | \text{wet}) = 0.25$
0.1/0.4

$P(RB_2 | \text{wet}) = 0.50$
0.2/0.4

$P(RB_3 | \text{wet}) = 0.25$
0.1/(0.4)

$\Sigma = 1.00$

DISCRETE RV

- Conditional Distribution $\qquad p_{X|Y}(x|y) = \dfrac{p_{XY}(x,y)}{p_Y(y)}$
- Joint Distribution $\qquad p_{XY}(x,y) = p_Y(y) * p_{X|Y}(x|y)$

  Marginal * Conditional
- Marginal Distribution $\qquad p_X(x) = \sum_y [p_Y(y) * p_{X|Y}(x|y)]$

EXAMPLE

For example, the following matrix displays the probabilities of different weather conditions and of different recreation benefits levels obtained from use of a reservoir in a state park

$$\boxed{E[X]=\sum_i x_i p_X(x_i)}$$

Marg

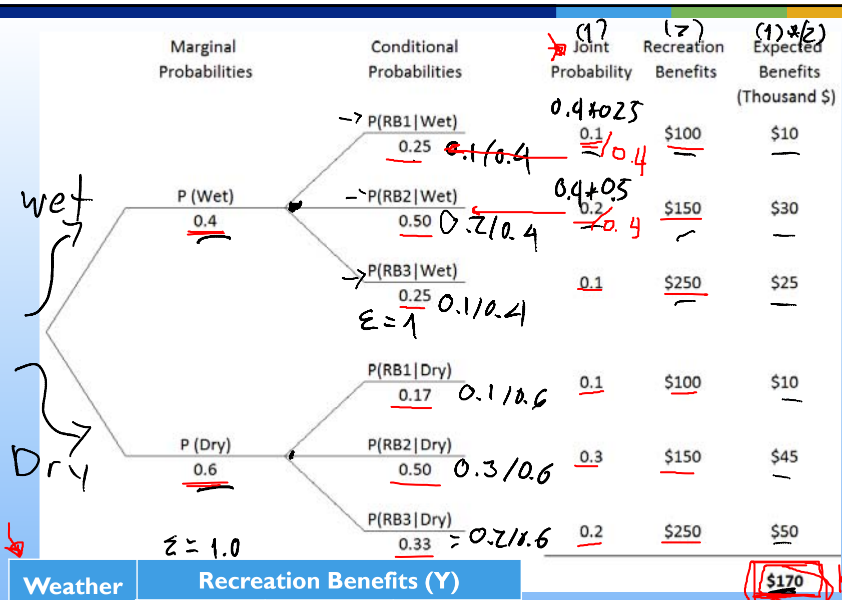
Expected Benefits

Weather (X)	Recreation Benefits (Y)			
	RB_1	RB_2	RB_3	
Wet	0.1	0.2	0.1	0.4
Dry	0.1	0.3	0.2	0.6

Σ=1.0

$$\overline{B} = 0.10 * RB_1 + 0.20 RB_2 + 0.10 RB_3 + 0.10 RB_1 + 0.30 RB_2 + 0.20 RB_3$$
$$= 0.10(\$100K) + 0.20(\$150K) + 0.10(\$250K)$$
$$+ 0.10(\$100K) + 0.30(\$150K) + 0.20(\$250K)$$
$$= \$170K$$

RB_1= \$100K
RB_2= \$150K
RB_3= \$250K



Weather (X)	Recreation Benefits (Y)			
	RB_1	RB_2	RB_3	
Wet	0.1/0.4 +	0.2 +	0.1	= 0.4
Dry	0.1 +	0.3 +	0.2	= 0.6

$$E[X]=\sum_i x_i p_X(x_i)$$

QUANTILES

The *pth quantile of a random variable X is the smallest* value x_p *such that X has a probability p of assuming a* value equal to or less than x_p

$$\Pr\{X < x_p\} \le p \le \Pr\{X \le x_p\}$$

$$\text{Median}: \ x_{0.50}$$

$$\text{Interquartile Range}: [x_{0.25}, x_{0.75}]$$

FLOW DURATION CURVE

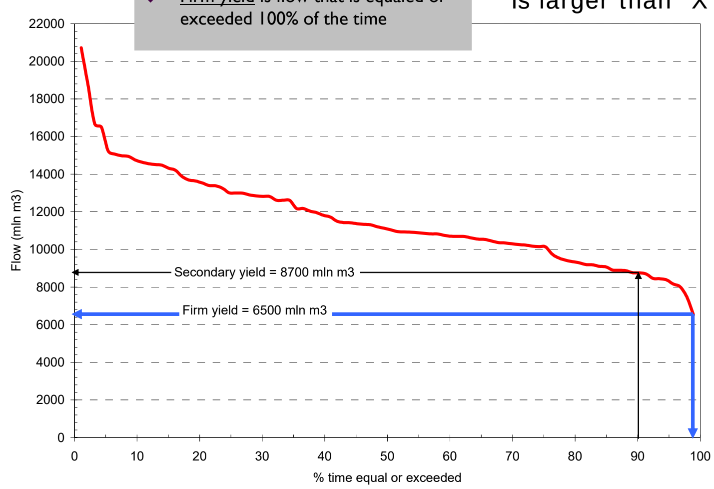
$$1 - p = \Pr\{X > x_p\} = 1 - \frac{i}{n+1}$$

Year	Flow	Rank	$P(X>x)=$	Ranked
			$1\text{-}i/(n+1)=$	Flow
	x	i	$1\text{-}p$	$x_{(i)}$
1911	10817	1	0.99	6525
1912	11126	2	0.98	7478
1913	11503	3	0.97	8014
1914	11428	4	0.96	8161
1915	10233	5	0.95	8378
…				
1997	10343	87	0.06	15062
1998	14511	88	0.05	15242
1999	14557	89	0.04	16504
2000	12614	90	0.03	16675
2001	12615	91	0.02	18754
2002	16675	92=n	0.01	20725

FLOW DURATION CURVE

- Flow duration curve - Discharge vs % of time flow is equaled or exceeded.
- Firm yield is flow that is equaled or exceeded 100% of the time

P(X>x)
Probability that flows is larger than "X"



UNITS

- 1 ft = 0.3048 m
- 1 m^3 = 28.3168x10 $^{-3}$ ft^3
- 1 m^3 = 35.3147 ft^3
- 1 ha = 10,000 m^2
- 1 acre = 43,560 ft^2
  = 0.4047 ha
  = 4047 m^2
- 1 gal = 3.785x10 $^{-3}$ m^3
  = 3.785 L

- 1 m^3 = 8.11x10^{-4} af
  10^9 m^3 = 8.11x10^5 af
  1 km^3 = 0.811 maf
- 1 m^3 = 264 gal
  10^9 m^3 = 264x10^9 gal
  1 km^3 = 264 bg
  1 km^3/yr = 0.7234 bgd