Expected Monetary Value

ESM-121 Water Science and Management

Samuel Sandoval Solis, PhD
Assistant Professor

Presentation 4 of 10

**BASIC CONCEPTS**

- **The mean**

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

- **Weighted mean**

  \[ \bar{X} = \frac{\sum_{i=1}^{n} X_i \cdot W_i}{\sum_{i=1}^{n} W_i} \]

- **Median**

  \[
  \text{median} (P_{a,b}) = \begin{cases} 
  X_{(n+1)/2} & \text{when } n \text{ is odd, and} \\
  \frac{1}{2} (X_{(n/2)} + X_{(n/2)+1}) & \text{when } n \text{ is even.}
  \end{cases}
  \]
Figure 1.1  Density Function for a Lognormal Distribution

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:

(a) 2 4 8 9 11 11 12  
\[ \bar{X} = 8.1 \]
\[ P_{.50} = 2 \]

(b) 2 4 8 9 11 11 11 12
\[ \bar{X} = 23.6 \]
\[ P_{.50} = 2 \]
Variance and Standard Deviation

\[ s^2 = \frac{\sum (X_i - \bar{X})^2}{n-1} \]

\[ s = \sqrt{s^2} \]

Interquartile Range

\[ \text{IQR} = P_{0.75} - P_{0.25} \]

Median Absolute Deviation

\[ \text{MAD}(X_i) = \text{median} |d_i|, \quad \text{where } d_i = X_i - \text{median}(X_i) \]
### STD. DEVIATION VS IQR VS MAD

<table>
<thead>
<tr>
<th>data</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>9</th>
<th>11</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(x_i - \bar{x})^2$</td>
<td>37.2</td>
<td>16.8</td>
<td>0.01</td>
<td>0.81</td>
<td>8.41</td>
<td>8.41</td>
<td>15.2</td>
</tr>
<tr>
<td>$</td>
<td>d_i</td>
<td>=</td>
<td>x_i - P_{50}</td>
<td>$</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

IQR = $11 - 4 = 7$

$s^2 = (3.8)^2$

MAD = median $|d_i| = 2$

<table>
<thead>
<tr>
<th>data</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>9</th>
<th>11</th>
<th>11</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(x_i - \bar{x})^2$</td>
<td>37.2</td>
<td>16.8</td>
<td>0.01</td>
<td>0.81</td>
<td>8.41</td>
<td>8.41</td>
<td>12522</td>
</tr>
<tr>
<td>$</td>
<td>d_i</td>
<td>=</td>
<td>x_i - P_{50}</td>
<td>$</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

IQR = $11 - 4 = 7$

$s^2 = (42.7)^2$

MAD = median $|d_i| = 2$

---

### HISTOGRAMS

**Figure 2.2a.** Histogram of annual streamflow for the Letting River

**Figure 2.2b.** Second histogram of same data, but with different interval divisions
Figure 2.4: Quantile plot of the Licking River annual streamflow data.
Random Variables (RV) Probabilities Variables

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

- Discrete RV (\(\text{Discrete} \quad \text{RV}\))
  - Takes on values from a discrete set
    - \# of years until a certain flood stage returns
    - \# of times reservoir storage drops below a level

- Continuous RV (\(\text{Continuous} \quad \text{RV}\))
  - Takes on values from a continuous set
    - e.g., Rainfall, Streamflow, Temperature, Concentration

\[
f_X(x) = \frac{dF_X(x)}{dx}
\]

\[
P_X(x_i) = \Pr[X = x_i]
\]
Continuous RV

Discrete RV

\[ F_X(x) = \Pr\{X \leq x\} \]

\[ F_X(x) = \sum_{x_i \leq x} p_X(x_i) \]

Note (expected value of \( X \) – Mean of \( X \))

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

\[ E[X] = \sum_{i} p_X(x_i) \]

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

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

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

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

**Expected Benefits**

<table>
<thead>
<tr>
<th>Weather (X)</th>
<th>Recreation Benefits (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet</td>
<td>RB_1</td>
</tr>
<tr>
<td>Dry</td>
<td>RB_1</td>
</tr>
</tbody>
</table>

\[ \text{RB}_1 = $100K \]
\[ \text{RB}_2 = $150K \]
\[ \text{RB}_3 = $250K \]
The joint distribution of two RVs, $X$ and $Y$:

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

**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.

<table>
<thead>
<tr>
<th>Weather (X)</th>
<th>Recreation Benefits (Y)</th>
<th>$\sum x_i p_X(x_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet</td>
<td>RB₁: $100K$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RB₂: $150K$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RB₃: $250K$</td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td>RB₁: $0.1$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RB₂: $0.2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RB₃: $0.1$</td>
<td></td>
</tr>
</tbody>
</table>

$$E[X] = \sum x_i p_X(x_i)$$
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.

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

### Weather ($X$) vs. Recreation Benefits ($Y$)

<table>
<thead>
<tr>
<th>Weather (X)</th>
<th>Recreation Benefits (Y)</th>
<th>Joint Probability</th>
<th>Recreation Benefits</th>
<th>Expected Benefits (Thousand $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet</td>
<td>RB₁</td>
<td>0.1</td>
<td>$100</td>
<td>$10</td>
</tr>
<tr>
<td></td>
<td>RB₂</td>
<td>0.2</td>
<td>$150</td>
<td>$30</td>
</tr>
<tr>
<td></td>
<td>RB₃</td>
<td>0.1</td>
<td>$250</td>
<td>$25</td>
</tr>
<tr>
<td>Dry</td>
<td>RB₁</td>
<td>0.1</td>
<td>$100</td>
<td>$10</td>
</tr>
<tr>
<td></td>
<td>RB₂</td>
<td>0.3</td>
<td>$150</td>
<td>$45</td>
</tr>
<tr>
<td></td>
<td>RB₃</td>
<td>0.2</td>
<td>$250</td>
<td>$50</td>
</tr>
</tbody>
</table>

$$E[X] = \sum_i x_i \Pr_X(x_i)$$
Two RVs X and Y can have a joint distribution $F_{XY}(x,y)$.

The **marginal distribution** of X is the distribution of X ignoring Y.

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

<table>
<thead>
<tr>
<th>Weather (X)</th>
<th>Recreation Benefits (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet</td>
<td>RB1</td>
</tr>
<tr>
<td>Dry</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Weather (X) | Recreation Benefits (Y) |
-------------|--------------------------|
| Wet | RB1 | RB2 | RB3 |
| Dry | 0.1 | 0.2 | 0.1 |

$$E[X] = \sum_{i} x_i p_X(x_i)$$
**Conditional Distributions**

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

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

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

**Weather (X)** | **Recreation Benefits (Y)**
---|---|---
**Wet** | 0.1 | 0.2 | 0.1
**Dry** | 0.1 | 0.3 | 0.2

**Discrete RV**

- **Conditional Distribution**
  $$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

- **Joint Distribution**
  $$p_{X,Y}(x,y) = p_Y(y) \cdot p_{X|Y}(x|y)$$

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

\[
\begin{array}{ccc}
\text{Weather (X)} & \text{Recreation Benefits (Y)} \\
& \text{RB}_1 & \text{RB}_2 & \text{RB}_3 \\
\text{Wet} & 0.1 & 0.2 & 0.1 \\
\text{Dry} & 0.1 & 0.3 & 0.2 \\
\end{array}
\]

\[
\sum_{i} x_i p_X(x_i) = E[X] = \sum_{i} x_i p_X(x_i)
\]

**Expected Benefits**

\[
\begin{align*}
\bar{B} &= 0.10 \times \text{RB}_1 + 0.20 \times \text{RB}_2 + 0.10 \times \text{RB}_3 + 0.10 \times \text{RB}_1 + 0.30 \times \text{RB}_2 + 0.20 \times \text{RB}_3 \\
&= 0.10(\$100\text{K}) + 0.20(\$150\text{K}) + 0.10(\$250\text{K}) + 0.10(\$100\text{K}) + 0.30(\$150\text{K}) + 0.20(\$250\text{K}) \\
&= \$170\text{K}
\end{align*}
\]

**Marginal Probabilities**

\[
\begin{array}{c}
\text{Wet} \\
\text{Dry} \\
\end{array}
\]

\[
\begin{array}{c}
0.4 \\
0.6 \\
\end{array}
\]

**Conditional Probabilities**

\[
\begin{array}{c}
\text{Wet} \\
\text{Dry} \\
\end{array}
\]

\[
\begin{array}{c}
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\end{array}
\]

\[
\begin{array}{c}
0.25 \\
0.50 \\
0.25 \\
0.40 \\
0.30 \\
0.30 \\
0.20 \\
0.17 \\
0.33 \\
\end{array}
\]

\[
\begin{array}{c}
0.1 \\
0.1 \\
0.1 \\
0.4 \\
0.3 \\
0.2 \\
0.4 \\
0.6 \\
0.6 \\
\end{array}
\]

\[
\begin{array}{c}
\text{Wet} \\
\text{Dry} \\
\text{Wet} \\
\text{Wet} \\
\text{Dry} \\
\text{Wet} \\
\text{Dry} \\
\text{Dry} \\
\text{Dry} \\
\end{array}
\]

\[
\begin{array}{c}
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\end{array}
\]

\[
\begin{array}{c}
\text{Wet} \\
\text{Dry} \\
\text{Wet} \\
\text{Wet} \\
\text{Dry} \\
\text{Wet} \\
\text{Dry} \\
\text{Dry} \\
\text{Dry} \\
\end{array}
\]

\[
\begin{array}{c}
\$100 \\
\$150 \\
\$250 \\
\$100 \\
\$150 \\
\$250 \\
\$100 \\
\$150 \\
\$250 \\
\$150 \\
\end{array}
\]

\[
\begin{array}{c}
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\text{RB}_1 \\
\text{RB}_2 \\
\text{RB}_3 \\
\end{array}
\]

\[
\begin{array}{c}
\$0 \\
\$0 \\
\$0 \\
\$0 \\
\$0 \\
\$0 \\
\$0 \\
\$0 \\
\$0 \\
\end{array}
\]

\[
\begin{array}{c}
\$10 \\
\$30 \\
\$25 \\
\$10 \\
\$45 \\
\$50 \\
\$70 \\
\end{array}
\]

\[
\begin{array}{c}
\text{E}[X] = \sum_{i} x_i p_X(x_i)
\end{array}
\]
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 $x_p$.

$$\mathbb{P}(X < x_p) \leq p \leq \mathbb{P}(X \leq x_p)$$

Median: $x_{0.50}$

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

### Flow Duration Curve

$$1 - p = \mathbb{P}(X > x_p) = 1 - \frac{i}{n+1}$$

<table>
<thead>
<tr>
<th>Year</th>
<th>Flow</th>
<th>Rank</th>
<th>$1 - i/(n+1)$</th>
<th>$\mathbb{P}(X &gt; x) = 1 - \frac{i}{n+1}$</th>
<th>Ranked Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>1911</td>
<td>10817</td>
<td>1</td>
<td>0.99</td>
<td>$1 - \frac{1}{10+1} = 0.99$</td>
<td>6525</td>
</tr>
<tr>
<td>1912</td>
<td>11126</td>
<td>2</td>
<td>0.98</td>
<td>$1 - \frac{2}{10+1} = 0.98$</td>
<td>7478</td>
</tr>
<tr>
<td>1913</td>
<td>11503</td>
<td>3</td>
<td>0.97</td>
<td>$1 - \frac{3}{10+1} = 0.97$</td>
<td>8014</td>
</tr>
<tr>
<td>1914</td>
<td>11428</td>
<td>4</td>
<td>0.96</td>
<td>$1 - \frac{4}{10+1} = 0.96$</td>
<td>8161</td>
</tr>
<tr>
<td>1915</td>
<td>10233</td>
<td>5</td>
<td>0.95</td>
<td>$1 - \frac{5}{10+1} = 0.95$</td>
<td>8378</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1997</td>
<td>10343</td>
<td>87</td>
<td>0.06</td>
<td>$1 - \frac{87}{92+1} = 0.06$</td>
<td>15062</td>
</tr>
<tr>
<td>1998</td>
<td>14511</td>
<td>88</td>
<td>0.05</td>
<td>$1 - \frac{88}{92+1} = 0.05$</td>
<td>15242</td>
</tr>
<tr>
<td>1999</td>
<td>14557</td>
<td>89</td>
<td>0.04</td>
<td>$1 - \frac{89}{92+1} = 0.04$</td>
<td>16504</td>
</tr>
<tr>
<td>2000</td>
<td>12614</td>
<td>90</td>
<td>0.03</td>
<td>$1 - \frac{90}{92+1} = 0.03$</td>
<td>16675</td>
</tr>
<tr>
<td>2001</td>
<td>12615</td>
<td>91</td>
<td>0.02</td>
<td>$1 - \frac{91}{92+1} = 0.02$</td>
<td>18754</td>
</tr>
<tr>
<td>2002</td>
<td>16675</td>
<td>92</td>
<td>0.01</td>
<td>$1 - \frac{92}{92+1} = 0.01$</td>
<td>20725</td>
</tr>
</tbody>
</table>
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”

1 ft = 0.3048 m
1 m³ = 28.3168x$10^{-3}$ ft³
1 m³ = 35.3147 ft³
1 ha = 10,000 m²
1 acre = 43,560 ft²
  = 0.4047 ha
  = 4047 m²
1 gal = 3.785x$10^{-3}$ m³
  = 3.785 L

1 m³ = 8.11x$10^{-4}$ af
10⁹ m³ = 8.11x$10^5$ af
1 km³ = 0.811 maf

1 m³ = 264 gal
10⁹ m³ = 264x$10^9$ gal
1 km³ = 264 bg
1 km³/yr = 0.7234 bgd