[redacted] [redacted]

Chapter 2 [redacted]

Flood Frequency Analysis

~~Statistics~~
- Probability
- (~~Inferential~~) Statistics

Probability - estimating the chance of some event occurring.

(~~Inferential~~) Statistics - estimating parameters of distributions of results from an experiment.

A. Descriptive Statistics - parameters that describe a sample and only apply to the given sample.

B. Inferential Statistics - infer something about the total population.

[redacted]

PROBABILITY DISTRIBUTIONS

1) Frequency Histogram (Discrete variable)

Relative Frequency, $p = \frac{f}{N}$

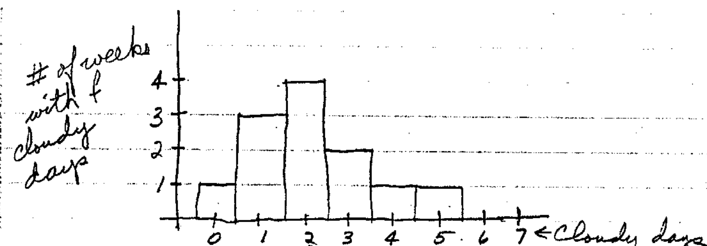
f = number of occurrence of the wanted event

N = total number of events (population)

Example:

Var(x) = Number of cloudy days per week at UNO

Observed for 12 weeks = N

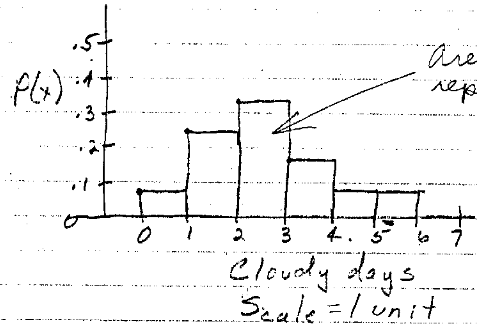


Cloudy days	Weeks	P(x)
0	1	1/12
1	3	1/4
2	4	1/3
3	2	1/6
4	1	1/12
5	1	1/12
6	0	0
7	0	0

Convert to Probabilitie's

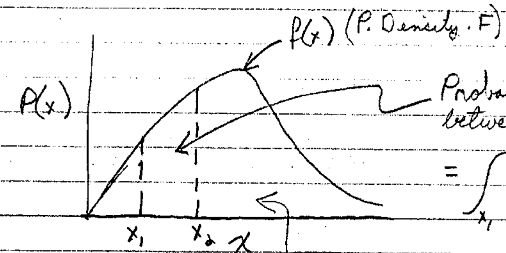
[redacted]

Probability Mass Function (for Discrete Random Variable)



area under the bars represents the probabilities

Probability Density Function (for continuous Random Variables)



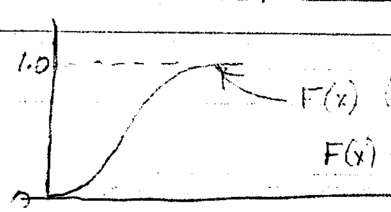
Probability of x between x_1 and x_2 $= \int_{x_1}^{x_2} f(x)\,dx$

Total area under curve = 1.

Probability of exceedance $F(x_1) = \int_{x_1}^{\infty} f(x)\,dx$

Cumulative Distribution Function

$\Sigma P(x)$

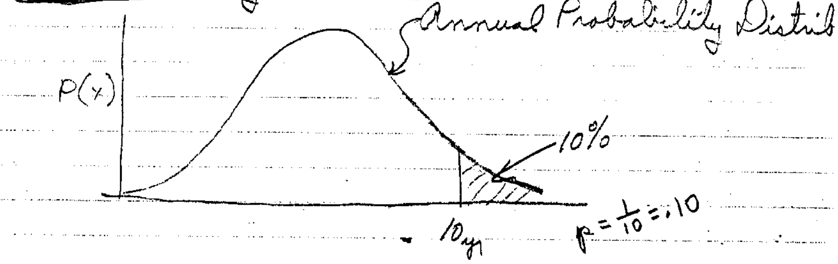


$F(x)$ (Cumulative Dist. F)

$F(x) = \int_0^x f(x)\,dx$

[redacted]

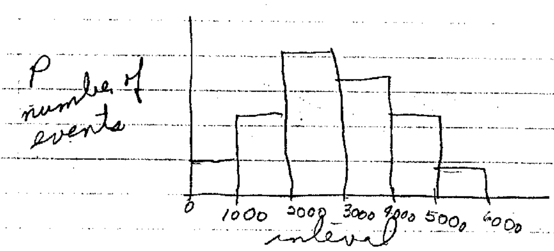
Prob. Density Function



Obtaining P D F from observed events.

1) Group magnitudes of flood together, i.e, 1000 cfs, 2000 cfs, etc.

2) Plot number of occurrences for each interval vs the interval



3) fit curve to observed data (least squares, etc.)

But we use the method of moments of Distribution to fit the curve.

[redacted]

Moments of Distribution

$$M_r^0 = \sum_{i=1}^{n} x_i^r f(x_i)$$ r th moment about the origin

where x_i = point

$f(x)$ = prob. of that point

measures central tendency

1st moment about "origin"

$r=1$

$$M = \sum_{i=1}^{n} x_i f(x_i) = \left\{ \begin{array}{l} \text{theoretical value} \\ \text{Expected Value} \\ \text{of Mean} = \mu \end{array} \right.$$

$$\overline{X} = \frac{\sum x_i}{n}$$ {estimate of μ from a sample

measures the dispersion about the mean

Second moment about "mean"

$$M_2^{\mu} = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) = \left\{ \begin{array}{l} \text{theoretical value} \\ \text{Variance} = \sigma^2 \end{array} \right.$$

$$s^2 = \frac{\sum (x_i - \overline{x})^2}{n-1}$$ {estimate of σ^2 from a sample

[redacted]

measures the symmetry about the mean

Third moment about "mean"

$$M_3 = \sum_{i=1}^{n} (x_i - \mu)^3 f(x_i) = \text{skewness} = \alpha$$

$$a = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (x - \bar{x})^3$$ sample skewness

Discriptive Statistics
Computational Formula's

Mean

$$\boxed{\bar{X} = \frac{\sum x}{n}}$$

Variance

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{n\left(\sum x_i^2\right) - \left(\sum x_i\right)^2}{n(n-1)} = \frac{\left(\sum x_i^2\right) - \frac{\left(\sum x_i\right)^2}{n}}{n-1}$$

Standard Deviation

$$s = \sqrt{s^2}$$

Coefficient of Variation

$$C_v = \frac{s}{X}$$

[redacted]

Skewness

$$a = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (x_i - \bar{x})^3$$

Coefficient of skewness

$$C_s \text{ or } g = \frac{a}{s^3}$$

$$\boxed{g = \frac{n^2 \left(\Sigma x_i^3\right) - 3n \left(\Sigma x_i\right)\left(\Sigma x_i^2\right) + 2\left(\Sigma x_i\right)^3}{n(n-1)(n-2)\, s^3}}$$

Note: for a good measure of skewness (and g) you need <u>at least 120 years of record.</u>

$$a = \frac{n^2 \left(\Sigma x_i^3\right) - 3n \left(\Sigma x_i\right)\left(\Sigma x_i^2\right) + 2\left(\Sigma x_i\right)^3}{n(n-1)(n-2)}$$

[redacted]

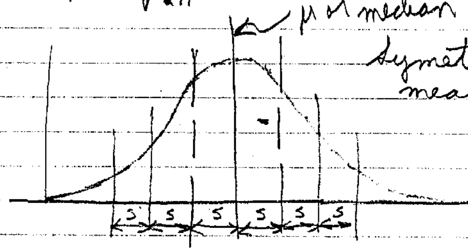
Distributions Used in Hydrology

Normal

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

two parameter distribution

μ or median



Symmetrical about mean therefore

$\alpha = 0$ (skewness)

68% of area is between $\pm s$

95% of area is between $\pm 2s$

99.7% of area is between $\pm 3s$

This does not occur very often in Hydrology.

[redacted]

[redacted]

Log normal

Substitute $\ln x$ for x in Normal to get Log normal.

If you substract "a" (skewness) from each x in series, we get the 3 parameter log normal.

Gamma

$$f(x) = \frac{x^{\alpha} e^{-x/\beta}}{\beta^{\alpha+1} \Gamma(\alpha+1)}$$

2 parameter Gamma Distribution

11/16/87

Pearson

3 parameter gamma distribution. This is the function most used in Hydrology.

* Log Pearson Type III

$$f(x) = \frac{|\alpha|}{\Gamma(\lambda)} \frac{e^{\alpha n}}{x^{1+\alpha k}} \left[\alpha (k \ln x - n)\right]^{\lambda-1}$$

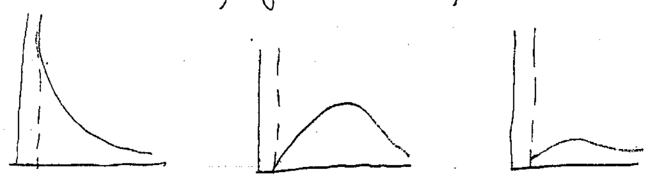
where α, λ, n parameters

$k = \log_{10} e$

Note: On page 765 of text is the Pearson Type III therefore must take log of data to use.

[redacted]

when $\alpha > 0$, pos skew, has lower bound



when $\alpha < 0$, negatively skewed, distribution has an upper bound which is not what we want since floods do not have an upper bound.

Gumbel Distribution

Cumulative Distribution

$$F(x) = e^{-e^{(-\alpha(x-\mu))}}$$