Time Series Analysis

AST5003F, 2020
Lecture #2

this lecture

- Recap of stats
- Moving statistics using pandas
- Smoothing data
- Resampling data
- Trends
    - Using filters, using polynomials
    - Seasonality (if a known cycle, i.e. annual, or if orbital period)

Recap of some stats

- From the baseline assessment, we will recap some introductory statistics.

Populations & samples

Statistics to describe a population of M members

Measures of the central value:

Mean $$\mu = \frac{1}{M}\sum_{i=1}^{M} x_i$$

Median $$n(x_i \leq \mu_{1/2}) = n(x_i \geq \mu_{1/2}) \sim \frac{M}{2}$$

Mode $$n(x_i = \mu_{\max}) > n(x_i = y, y \neq \mu_{\max})$$

Populations & samples

Statistics to describe a population of M members

Table 2.3. *Employee salaries at astroploitcom*

Job title (number of employees)	Salary in thousands of dollars
President (1)	2000
Vice president (1)	500
Programmer (3)	30
Astronomer (4)	15

Mean	Median	Mode
300	30	15

Populations & samples

Statistics to describe a population of M members

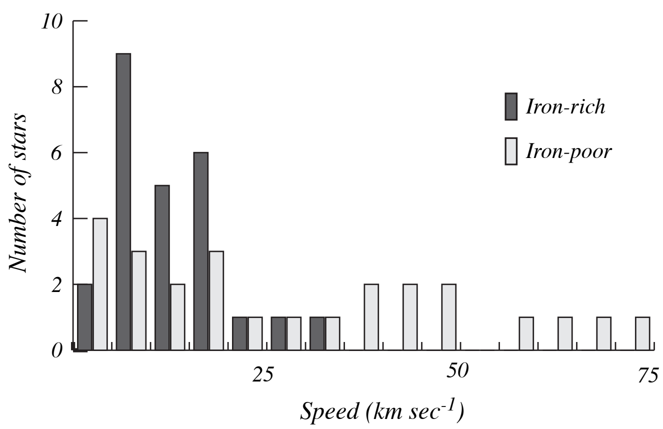
Measures of dispersion:

Variance:

$$\sigma^2 = \frac{1}{M}\sum_{i=1}^{M}(x_i - \mu)^2$$

Standard deviation:

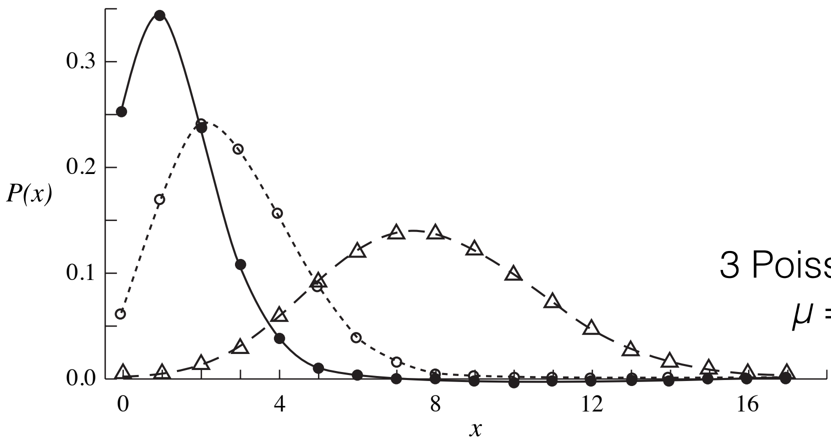
$$\sigma = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(x_i - \mu)^2}$$



Poisson distribution

$$P_p(x, \mu) = \frac{\mu^x}{x!} e^{-\mu}$$

- Discrete distribution
- Good for counting experiments
- Probability of getting x events in some given time interval



3 Poisson distributions
μ = 1.4, 2.8, 8

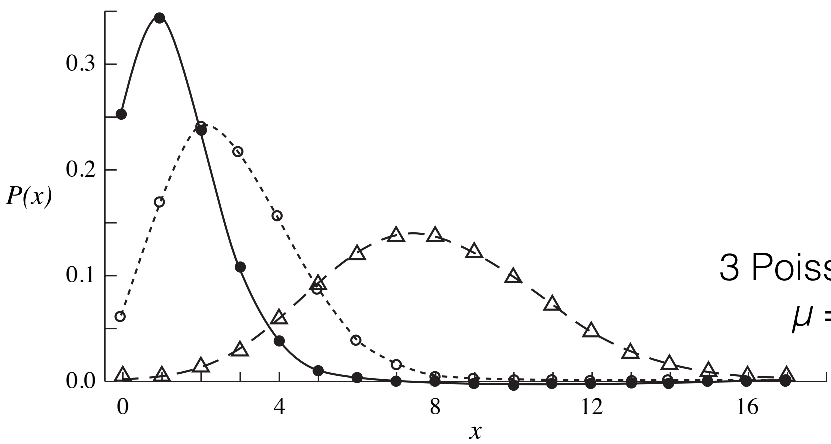
Poisson distribution

As μ increases, so does the variance of the distribution.

$$\sigma^2 = \mu$$

Fractional uncertainty in counting N events:

$$\frac{\sigma}{\mu} = \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$$



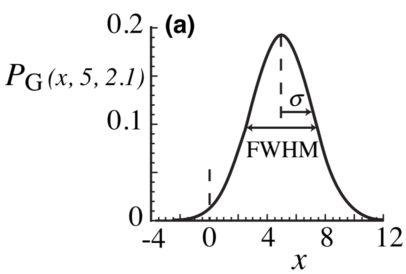
3 Poisson distributions
μ = 1.4, 2.8, 8

Gaussian (normal) distribution

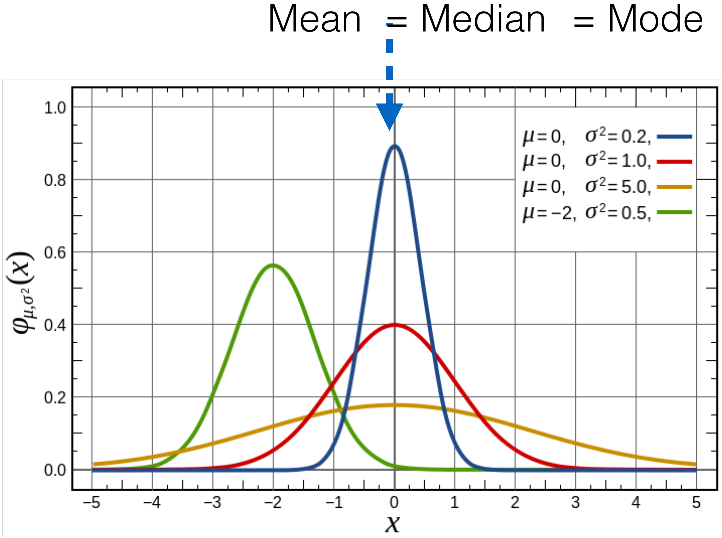
$$P_G(x, \mu, \sigma)dx = \frac{dx}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

- Bell curve
- μ and σ independent of each other

Mean = Median = Mode

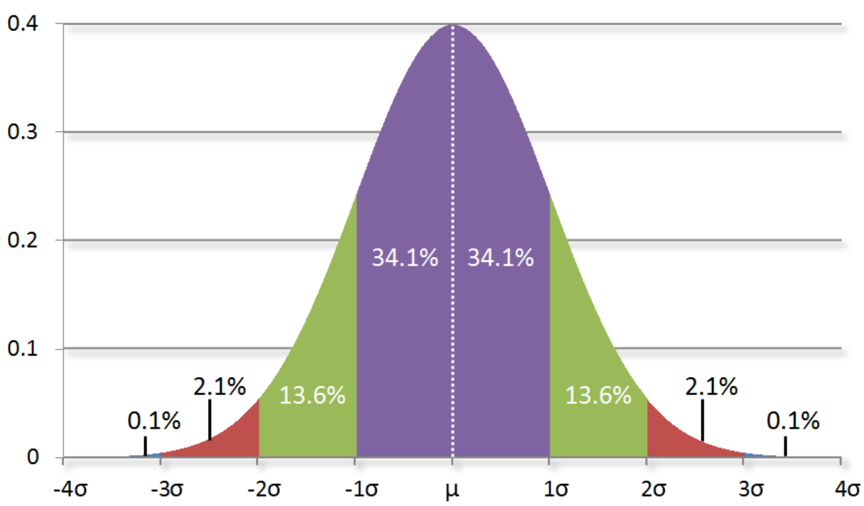


FWHM = 2.354 σ



Gaussian distribution & confidence intervals

- Approx 2/3 of the time a measurement will fall within *1σ* of the mean
- 95% of the time, a measurement will fall within *2σ* of the mean
- Only 3 in 1000 measurements (99.7%) will fall outside of *3σ* of the mean



Estimating uncertainty

How do you estimate the uncertainty of a particular quantitative measurement?

Central Limit theorem:

If $\{x_1, x_2, \ldots, x_n\}$ is a sequence of n independent variables drawn from $P(\mu,\sigma)$, then as n becomes large, the distribution of the variables

$$\bar{x_n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

will approach a Gaussian with mean μ and variance σ^2/n.

$$\sigma_\mu(n) = \frac{s}{\sqrt{n}}$$

Standard deviation of the mean