NOTES FROM PHYS 600 CLASS BASED ON FEIGELSON and BABU BOOK

The following summaries were compiled after having gone through the Statistical Astronomy class in 2013, 2015, and 2018.

1. Preliminaries

1.1. Definitions

Probability distribution function: p(x)

Sample distribution function: f(x)

Mean: Expected value of x = E(x) = $\mu = \int_{-\infty}^{\infty} xp(x)dx$

Variance: E(x^2) − E^2(x) = E(x−μ)^2 = σ^2

Moment Generating Function: E(e^{tx}). Expand this in powers of t to get moments of the distribution, like E(x), E(x^2), etc.

Standard deviation: σ

Empirical Distribution Function: e.d.f. = $\widehat{F}(x) = \int_{-\infty}^{x} f(x)dx$

Cumulative Distribution Function: c.d.f. = F(x) = $\int_{-\infty}^{x} p(x)dx$

Trimmed Mean: Like a mean but you reject a certain number of high and low points. IRAF does the same thing with the imcombine task. With the maximum rejections you end up with the median.

Interquartile Range: Where 25% and 75% of the data fall. Indicated with the box part of a box plot.

Box-whiskers: Box usually defined as above. No consensus on whiskers. Outlier points are plotted individually.

MAD: median absolute deviation = Med$|X_i - Med|$ is an excellent measure of dispersion in contaminated data.

Heteroscedastic/Homoscedastic: Errors or scatter vary/don't vary with the value of the variable.

1.2. Concepts

Robust: Results that are insensitive to outlier points. For example, the median is robust but the mean is not.

Breakdown: Fraction of data that need to be contaminated to wreck a statistic. For the median it is 50% .

Confidence Intervals: Typically used to reject at some level of confidence that some condition exists in the data. For example, if getting a correlation as good as observed occurs less than 1% of the time for data that are really uncorrelated so the apparent correlation is caused by noise, then we reject with 99% confidence the hypothesis that the data are uncorrelated. Another example: you can reject the hypothesis that the mean of the data $\mu > 100$ with 99% certainty.

1.3. Probability Distribution Functions

Binomial: Like a series of coin flips, this distribution is the probability of getting x successes in n attempts when the probability of a success is θ for each attempt. The mean $\mu = \theta$ and the variance $\sigma^2 = \theta(1-\theta)$.

$$p(x) = \frac{n!}{x!(n-x)!}\theta^x(1-\theta)^{n-x}$$

Poisson: Limit of the binomial when n $\to \infty$, $\theta \to 0$ such that nθ=λ is a constant. Also occurs in counting statistics where the chance of success in a time interval is λ, each count is independent of the previous history, and the probability of more than one count in a time interval is negligible. This is the distribution of shot noise, where the signal-to-noise ratio $\mu/\sigma = \lambda^{1/2} = \mu^{1/2}$, so S/N $\sim$ C$^{1/2}$ where C are the counts. $\mu = \sigma^2 = \lambda$.

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Normal: The usual bell-curve. The half-width of the bell curve is something close to the standard deviation. A standard Normal has $\mu = 0$ and $\sigma = 1$. The distribution of $\bar{x}$ is normal. That is, if you repeatedly sample a population and measure $\bar{x}$ and plot what you get, it will look like a Normal distribution. Note that x doesn't need to be distributed as a Normal for $\bar{x}$ to be Normal.

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

Pareto: Fancy name for a power law, but one that is correctly normalized.

Gamma: A general distribution that has many well-known specific examples. It is a combination of a power law and an exponential. The general form is

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

where Γ is the gamma function, a generalization of the notion of an integer factorial. The mean and variance are $\mu = = \beta\alpha$, and $\sigma^2 = \beta^2\alpha$. Reduces to the *exponential distribution* for $\alpha = 1$ and $\beta = \theta$, and to the *Chi-square distribution* when $\alpha = \mathrm{n}/2$ and $\beta = 2$.