Some Notes on Statistics

Carl. W. Akerlof
January 7, 2008

Probability Distribution Functions:

$$\frac{dp}{dx} = f(x) \; ; \; \int_{x_{\min}}^{x_{\max}} f(x)dx = 1$$

Cumulative distribution function:

$$F(x) = \int_{x_{\min}}^{x} f(x')dx'$$

$$F(x_{\min}) = 0 \; ; \; F(x_{\max}) = 1$$

Population statistics:

Mean: $\langle x \rangle = \int_{x_{\min}}^{x_{\max}} x\, f(x)\, dx$

Variance: $\sigma_{\mathrm{x}}^2 = \langle x^2 \rangle - \langle x \rangle^2 = \int_{x_{\min}}^{x_{\max}} x^2 \; f(x)dx - \langle x \rangle^2$

Standard Deviation: $\sigma_x = \sqrt{\sigma_x^2}$

Sum of variables: $x_t = \sum_{i=1}^{n} x_i$

Mean: $\langle x_t \rangle = \sum_{i=1}^{n} \langle x_i \rangle$

Variance: $\sigma_{x_t}^2 = \sum_{i=1}^{n} \sigma_{x_i}^2$

Sample Statistics:

Sample Mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Sample Variance: $\sigma_s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i \right)^2 \right)$

Sample Standard Deviation: $\sigma_s = \sqrt{\sigma_s^2}$

Mean of Sample Means: $\langle \bar{x} \rangle = \langle x \rangle$

Mean of Sample Variances: $\langle \sigma_s^2 \rangle = \sigma_x^2$

Variance of Sample Means: $\sigma_{\bar{x}}^2 = \frac{1}{n}\sigma_x^2$

Variance of Sample Variances:

$$\sigma^2_{\sigma_s^2} = \frac{1}{n}\langle x^4\rangle - \frac{4}{n}\langle x^3\rangle\langle x\rangle - \frac{n-3}{n(n-1)}\langle x^2\rangle^2 + \frac{4(2n-3)}{n(n-1)}\langle x^2\rangle\langle x\rangle^2 - \frac{2(2n-3)}{n(n-1)}\langle x\rangle^4$$

For Gaussian distributions; $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$

Variance of sample variances: $\sigma^2_{\sigma_s^2} = \frac{2\sigma^4}{n-1}$

Standard deviation of sample variance: $\sigma_{\sigma_s^2} = \sqrt{\frac{2}{n-1}}\ \sigma^2$

Variance of sample standard deviations: $\sigma^2_{\sigma_s} = \frac{\sigma^2}{2(n-1)}$

Standard deviation of sample standard deviations: $\sigma_{\sigma_s} = \frac{\sigma}{\sqrt{2(n-1)}}$

Variance of median: $\sigma^2_{x_{1/2}} = \frac{1}{4nf^2\left(x_{1/2}\right)}$; $F(x_{1/2}) \equiv \frac{1}{2}$

For Gaussian distribution: $\sigma^2_{x_{1/2}} = \frac{\pi}{2n}\sigma^2$

Robust estimate of dispersion: Inter-Quartile Difference (IQD)

Assume $F(x_{1/4}) \equiv \frac{1}{4};\ F(x_{3/4}) \equiv \frac{3}{4}$

Then: $\text{IQD} = x_{3/4} - x_{1/4}$

For a Gaussian distribution:

$$\frac{\text{IQD}}{q^*} = \sigma \quad \text{where} \quad q^* = 1.3489798\ ;\ \frac{1}{2} = \frac{1}{\sqrt{2\pi}}\int_{-q^*/2}^{+q^*/2} e^{-\frac{x^2}{2}}dx$$

$$\sigma^2_{(\text{IQD}/q^*)^2} = \frac{2\pi}{n}\ \frac{e^{\left(\frac{q^*}{2}\right)^2}}{q^{*2}}\sigma^4$$

$$= 2.7209\frac{2\sigma^4}{n}$$