Probability - Chapter 5 - Statistics.

Sample Mean:

Let $X_1, X_2, \ldots, X_n$ be independent random variables of random sample of size n.

$$\hat{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

If $x_1, x_2, \ldots, x_n$ denote values obtained from sample size n:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Distribution of means:

Mean of sampling distributions of means $\mu_{\bar{x}}$ is

$$E(\bar{X}) = \mu_{\bar{x}} = \mu.$$

Theorem:

If population is infinite and sampling with replacement, variance $\sigma_{\bar{X}}^2$ is

$$E[(\bar{X} - \mu)^2] = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}.$$

Theorem:

If population size N and sampling without replacement, if sample size $n \leq N$:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right).$$

Standardized Variable:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$ is asymptotically normal i.e $\lim_{n \to \infty} P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} \, du.$

Sampling Distribution of Proportions:

With μ_p mean and σ_p standard distribution:

$$\mu_p = p, \quad \sigma_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}.$$

Sampling Differences and Sums:

$$\mu_{S_1 - S_2} = \mu_{S_1} - \mu_{S_2}, \quad \sigma_{S_1 - S_2} = \sqrt{\sigma_{S_1}^2 + \sigma_{S_2}^2}.$$

Sample means:

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2, \quad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_{\bar{X}_1}^2 - \sigma_{\bar{X}_2}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Standardized Variable:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

Binomially Distributed:

$$\mu_{P_1 - P_2} = \mu_{P_1} - \mu_{P_2} = P_1 - P_2, \quad \sigma_{P_1 - P_2} = \sqrt{\sigma_{P_1}^2 + \sigma_{P_2}^2} = \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}.$$

Sample Variance:

$$S^2 = \frac{(X_1-X)^2 + \dots + (X_n-X)^2}{n} \qquad (*)$$

$$E(S^2) = \mu_{S^2} = \frac{n-1}{n}\sigma^2$$

If sampling without replacement from a finite population:

$$E(S^2) = \mu_{S^2} = \left(\frac{N}{N-1}\right)\left(\frac{n-1}{n}\right)\sigma^2$$, as $n \to \infty$ this reduces to $(*)$.

Population Variance Unknown:

Let $T = \dfrac{\bar{X}-\mu}{\hat{S}/\sqrt{n}} = \dfrac{\bar{X}-\mu}{S/\sqrt{n-1}}$, $n-1$ degrees of freedom.

Theorem:

Random variables m, n from populations with variances σ_1^2, σ_2^2.

Then if the variances of random samples are given by S_1^2, S_2^2, the static

$$F = \frac{mS_1^2/(m-1)\sigma_1^2}{nS_2^2/(n-1)\sigma_2^2} = \frac{\hat{S}_1^2/\sigma_1^2}{\hat{S}_2^2/\sigma_2^2}$$, $m-1$ and $n-1$ degrees of freedom.

Grouped Data:

Frequency $n = f_1 + \dots + f_k = \sum f$

$$\bar{x} = \frac{f_1x_1 + \dots + f_kx_k}{n} = \frac{\sum fx}{n}.$$

Variance $$s^2 = \frac{f_1(x_1-\bar{x})^2 + \dots + f_k(x_k-\bar{x})^2}{n} = \frac{\sum f(x-\bar{x})^2}{n}.$$

Transformation:

Let $x = a + cu$,

$$\bar{x} = a + \frac{c}{n}\sum fu = a + c\bar{u}$$

$$s^2 = c^2\left[\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2\right]$$

Moments:

$$m_r = \frac{f_1(x_1-\bar{x})^r + \dots + f_k(x_k-\bar{x})^r}{n} = \frac{\sum f(x-\bar{x})^r}{n}.$$

$$m_r' = \frac{f_1x_1^r + \dots + f_kx_k^r}{n} = \frac{\sum fx^r}{n}.$$