Sample variance: $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$. Equivalent alternative formula: $s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$

Sample z-score for the ith observation: $z_i = \frac{x_i - \bar{x}}{s}$

If we transform the data using the linear transformation $x^* = a + bx$, then:

$$\bar{x}^* = a + b\bar{x}, s_{x^*} = |b|s_x, s_{x^*}^2 = b^2 s_x^2$$

Probability

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

$P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$.

$P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Two events A and B are independent if and only if:

$P(A \cap B) = P(A) \cdot P(B), P(A|B) = P(A), P(B|A) = P(B)$.

The Expected Value and Variance of Discrete Random Variables

$E(X) = \mu = \sum xp(x)$.

$\sigma^2 = E[(X-\mu)^2] = \sum (x-\mu)^2 p(x)$.

A handy relationship: $E[(X-\mu)^2] = E(X^2) - [E(X)]^2$.

Properties of Expectation and Variance

$E(a + bX) = a + bE(X)$, $\sigma_{a+bX}^2 = b^2\sigma_X^2$, $\sigma_{a+bX} = |b|\sigma_X$

If X and Y are both random variables then $E(X+Y) = E(X) + E(Y)$ and $E(X-Y) = E(X) - E(Y)$.

If X and Y are independent: $\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$ and $\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$

Discrete Probablity Distributions

Binomial distribution: $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$. $\binom{n}{x} = \frac{n!}{x!(n-x)!}$. $\mu = np, \sigma^2 = np(1-p)$.

Hypergeometric distribution: $P(X = x) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}$. $\mu = n\frac{a}{N}$.

Poisson distribution: $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \lambda = \mu = \sigma^2$.

Geometric distribution: $P(X = x) = (1-p)^{x-1}p$. $\mu = \frac{1}{p}$, $\sigma^2 = \frac{1-p}{p^2}$.

Normal Distribution

If X is normally distributed with a mean of μ and standard deviation σ, then $Z = \dfrac{X - \mu}{\sigma}$ has the standard normal distribution.

If $\bar{X}$ is the mean of n independent observations from a normal distribution with mean μ and standard deviation σ, then $Z = \dfrac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ has the standard normal distribution.

Inference Procedures for Means

(When sampling from a normally distributed population)

Inference for μ

If σ is known:

Confidence interval for μ: $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{X}}$, where $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$

To test H_0: $\mu = \mu_0$: $Z = \dfrac{\bar{X} - \mu_0}{\sigma_{\bar{X}}}$

If σ is unknown:

Confidence interval for μ: $\bar{X} \pm t_{\alpha/2} SE(\bar{X})$, where $SE(\bar{X}) = \dfrac{s}{\sqrt{n}}$

To test H_0: $\mu = \mu_0$: $t = \dfrac{\bar{X} - \mu_0}{SE(\bar{X})}$