Final Exam formulas

Binomial Distribution: $p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \ E(X) = np, \ VAR(X) = np(1-p)$

Geometric Distribution: $p(x) = p(1-p)^{x-1}, \ E(X) = \frac{1}{p}, \ VAR(X) = \frac{1-p}{p^2}$

Negative binomial distribution:

$$p(x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r, E(X) = \frac{r}{p}, Var(X) = \frac{r(1-p)}{p^2}$$

Hypergeometric distribution:

$$p(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \ E(X) = n\frac{r}{N}, \ VAR(X) = n\frac{r}{N}\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Poisson distribution: $p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, E(X) = \lambda, Var(X) = \lambda$

Uniform Distribution:

$$f(x) = 1/(b-a) \quad when \quad a \le x \le b; \qquad F(x) = \frac{x-a}{b-a}$$

$$\mu = (a + b)/2 \ and \ Var(X) = (b - a)^2/12.$$

Exponential distribution:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} \\ 0, \ elsewhere \end{cases}, \ 0 \le x \le \infty \qquad \lambda > 0$$

$$Mean = \mu = 1/\lambda, \qquad Variance = \sigma^2 = 1/\lambda^2, \qquad F(X) = 1 - e^{-\lambda x}$$

<u>Normal distribution</u>

A random variable X is said to have a normal distribution if for parameters σ>0 and -∞ < μ < ∞, the density function of X is

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

Standard Normal distribution

If x is an observation from a normal distribution that has mean μ and standard deviation σ, the standardized value of x is $z = \frac{x-\mu}{\sigma}$ and the standardized variable z has the standard normal distribution.

Sample Mean: $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$

Sample Standard Deviation: $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}}$

CV (Coefficient of variation): $\frac{s}{\bar{x}} \times 100$

Confidence Interval for population mean: $CI = \mu \pm \sigma z_{\frac{\alpha}{2}}$

Confidence Interval for sampling distribution: $CI = \mu \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}$

Confidence Interval for sampling distribution: $CI = \mu \pm \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}}$ (when σ *is not known*)

$X \sim N(\mu, \sigma^2)$; $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$; $\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$

$\hat{p} = \frac{x}{n}$, $E(\hat{p}) = p$, $se(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$

$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, $E(\bar{x}) = \mu$, $se(\bar{x}) = \sqrt{\frac{\sigma^2}{n}}$

Hypothesis testing : $z_{cal} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

Hypothesis testing : $t_{cal} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ (when σ *is not known*), *df = n-1*

Hypothesis testing (Two-means matched/paired test): $t_{cal} = \frac{\bar{d} - D_0}{\frac{s_d}{\sqrt{n}}}$, *df = n-1*

Correlation : $r = \frac{s_{xy}}{s_x s_y}$

$s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$, $s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2)}{n-1}}$, $s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2)}{n-1}}$

Regression coefficients: $\widehat{\beta_1} = b_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$, $\widehat{\beta_0} = b_0 = \bar{y} - b_1 \bar{x}$