Commerce 2QA3 Midterm Exam Formula Sheet

Quantitative Data Description (percentile, std. dev., mean, coefficient of var., quartiles, z-score, etc.):

Q3: 75^{th} percentile

Q1: 25^{th} percentile

Range=Max-Min

$IQR = Q3 - Q1$

Outlier Rule-of-thumb: $y < Q1 - 1.5 * IQR$ or $y > Q3 + 1.5 * IQR$

$$\bar{y} = \frac{\sum y}{n}$$

$$\text{Standard Deviation} = \sqrt{\text{Variance}}$$

$$s = \sqrt{\frac{\sum (y - \bar{y})^2}{n - 1}}$$

$$z = \frac{y - \mu}{\sigma}$$ "population"

$$z = \frac{y - \bar{y}}{s}$$ "sample"

$$CV = \frac{s}{\bar{y}}$$

Association, correlation and regression:

$$r = \frac{\sum z_x z_y}{n - 1}$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n - 1) s_x s_y}$$

$$cov(x, y) = r s_x s_y$$

$\hat{y} = b_0 + b_1 x$ where $b_1 = r \frac{s_y}{s_x}$ and $b_0 = \bar{y} - b_1 \bar{x}$

$$e = y - \hat{y}$$

Commerce 2QA3 Midterm Exam Formula Sheet

$$s_e = \sqrt{\frac{\sum e^2}{n-2}}$$

$$\hat{z}_y = r z_x$$

Probability:

$$P(A) = \frac{\text{No. of outcomes in A}}{\text{Total no. of outcomes}}$$

$$P(A) = 1 - \mathrm{P}(A^C)$$

$$P(A \text{ or } B) = P(A \cup B)$$

$$P(A \text{ and } B) = P(A \cap B)$$

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$P(A \text{ and } B) = P(A|B)P(B) = P(B|A)P(A)$$

Discrete Random Variables:

$$E(X) = \mu = \sum xP(x)$$

$$Var(X) = \sigma^2 = \sum (x-\mu)^2 P(x)$$

$$SD(X) = \sqrt{Var(X)}$$

Binomial: $P(X = x) = {}_n\mathrm{C}_x p^x q^{n-x} = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$

$\mu = np \quad \sigma = \sqrt{npq}$

Continuous Random Variables

Uniform Distribution: $f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \le x \le b \\ 0, & \text{otherwise} \end{cases}$ $\qquad P(c \le X \le d) = \frac{d-c}{b-a}$

$$E(X) = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{12}$$

Commerce 2QA3 Midterm Exam Formula Sheet

Adding two normally distributed random variables:

$X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$ and $Z = X + Y$ then $Z \sim N(\mu_x + \mu_y, \sqrt{\sigma_x^2 + \sigma_y^2})$

$X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$ and $K = X - Y$ then $K \sim N(\mu_x - \mu_y, \sqrt{\sigma_x^2 + \sigma_y^2})$

Sampling Distributions:

If conditions are satisfied, the sampling distribution of a proportion, $\hat{p} = \frac{x}{n}$, is

Normally distributed with $\mu(\hat{p}) = p$ and $SD(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$

Success/Failure condition: $np \geq 10$ and $n(1-p) \geq 10$

If conditions are satisfied, the sampling distribution of a mean, $\bar{y}$, is Normally distributed with $\mu(\bar{y}) = \mu_y$ and $SD(\bar{y}) = \frac{\sigma}{\sqrt{n}}$

Large sample condition: $n \geq 30$.

$sd = \frac{sd}{\sqrt{n}}$

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$SE(\bar{y}) = \frac{s}{\sqrt{n}}$$

higher = wider range
= more ME

$\hat{p}$rop

Confidence intervals and tests for one population for proportions and means:

pop

$$z = \frac{\hat{p} - p}{SD(\hat{p})} \quad \text{or} \quad z = \frac{\hat{p} - p}{SE(\hat{p})}$$

$1 - \frac{\alpha}{2} = z_\alpha$

estimate ±ME

$z^* = z_{\alpha/2}$

Confidence interval: $\hat{p} \pm z^* \times SE(\hat{p})$

Confidence interval for the difference between two population proportions:

Commerce 2QA3 Midterm Exam Formula Sheet

$(\hat{p}_1 - \hat{p}_2) \pm z^* \times SE(\hat{p}_1 - \hat{p}_2)$, where $SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

$n = \left(\frac{z_{\alpha/2}}{ME}\right)^2 \hat{p}(1-\hat{p})$ as a conservative measure to find the minimum sample size the sample proportion is assumed to be 50%

prop of success

mean of $\boxed{\text{sample prop}}$:

if $n \uparrow$ = SE $\downarrow$

ME $\downarrow$

$p = \frac{x\%}{n}$