STAT 235 Formula Sheet

Eddie Guo

Introduction

random asn \ random sampling	Y	N
Y	both	causal inferences
N	population inferences	neither

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right]$$

relative frequency $= \dfrac{f}{\sum f}$ $\qquad$ % = relative frequency × 100%

IQR = Q_3 - Q_1

- Mild outliers 1.5 IQR below Q_1 or above Q_3 (inner fence)
- Extreme outliers 3 IQR below Q_1 or above Q_3 (outer fence)
- EXCLUDE median when finding quartiles for odd n

Probability: Two Set Operations

Complement law: $P(A') = 1 - P(A)$

$(A \cup B)' = A' \cap B'$ $\qquad$ $(A \cap B)' = A' \cup B'$

$P(A \cup B) = 1 - P(A' \cap B')$

Mult law: $P(A \cap B) = P(B) \times P(A|B) = P(A) \times P(B|A)$

Add law: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Disjoint/mutually exclusive: $P(A \cup B) = P(A) + P(B)$

Conditional prob: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Indep events: $P(A \cap B) = P(A) \times P(B)$ given $P(A|B) = P(A)$

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B') \\ &= P(B) \times P(A|B) + P(B) \times P(A|B') \end{aligned}$$

Probability: Three Set Operations

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) \\ - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ + P(A \cap B \cap C) \end{aligned}$$

$$\begin{aligned} P(A \cap B' \cap C') &= P(A) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C) \\ &= P(A \cup B \cup C) - P(B) - P(C) + P(B \cap C) \end{aligned}$$

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

Deck of Cards

- 52 cards total w/ 13 cards in each suit
- 26 black, 26 red
- Each rank has 4 cards (4/52 = 1/13)
- Face cards: J, Q, K $\qquad$ 12 face cards (12/52 = 3/13)
- Diamonds (R) < clubs (B) < hearts (R) < spades (B)
- Black/red and face are independent

Counting Theory

Expt with k steps, and step i has n_i outcomes:

$$\text{Total no. of outcomes} = \prod_{i=1}^{k} n_i$$

Permutations: $P_k^n = \dfrac{n!}{(n-k)!}$ (order impt)

$P_0^n = 1 \qquad P_1^n = n \qquad P_n^n = n!$

Combinations: $C_k^m = \dbinom{n}{k} = \dfrac{n!}{k!(n-k!)}$ (order not impt)

$\dbinom{n}{0} = \dbinom{n}{n} = 1 \qquad \dbinom{n}{1} = n$

Discrete Random Variables

Probability mass function (pmf)

1. $f(x_i) = P(X = x_i)$
2. $0 \le f(x_i) \le 1$
3. $\sum_{i=1}^{n} f(x_i) = 1$

Cumulative distribution function (cdf)

$F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i)$

1. $0 \le F(x) \le 1$
2. If $x \le y$, then $F(x) \le F(y)$
3. $P(a \le X \le b) = F(b) - F(a-1)$

Expectations

$$\mu = \mathbb{E}(X) = \sum_{i=1} x_i p_i$$

$$\sigma^2 = Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \sum_{i=1}(x-\mu)^2 p_i = \sum_{i=1} x_i^2 p_i - \mu^2$$

Continuous Random Variables

Probability density function (pdf)

1. $f(x) \ge 0$
2. $\int_{-\infty}^{\infty} f(x)\ dx = 1$ (generic limits here)
3. $P(a \le X \le b) = \int_a^b f(x)\ dx$
4. $P(X = x) = 0 \quad \Rightarrow \quad P(a < X < b) = P(a \le X \le b)$

Cumulative distribution function (cdf)

$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\ du \qquad f(x) = \dfrac{dF}{dx}$

1. $0 \le F(x) \le 1$
2. If $x \le y$, then $F(x) \le F(y)$
3. $P(a \le X \le b) = F(b) - F(a)$

Expectations

$$\mu = \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\ dx \qquad \tilde{x} = F(x) = \frac{1}{2}$$

$$\sigma^2 = Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \int_{-\infty}^{\infty} x^2 f(x)\ dx - \mu^2$$

If X has pdf $f(x)$, $\mathbb{E}[h(X)] = \int_{-\infty}^{\infty} h(x) f(x)\ dx$

Linear Combinations of Random Variables

If $Y = a_1X_1 + \cdots + a_nX_n + b$, then $\mathbb{E}(Y) = a_1\mathbb{E}(X_1) + a_2\mathbb{E}(X_2) + \cdots + a_n\mathbb{E}(X_n) + b$

If all X_i are indep, $V(Y) = a_1^2V(X_1) + a_2^2V(X_2) + \cdots + a_n^2V(X_n)$

Means

1. $\mathbb{E}(a) = a$
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$
3. $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
4. $\mathbb{E}(aX \pm bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$

Variances

1. $V(a) = 0$
2. $V(aX) = a^2V(X)$
3. $V(aX + b) = a^2V(X)$
4. $V(aX \pm bY) = a^2V(X) + b^2V(Y) \pm 2ab \operatorname{cov}(X, Y)$
   If independent, $\operatorname{cov}(X, Y) = 0$

If X_i are indep rvs w/ $\mathbb{E}(X_i) = \mu$ and $V(X_i) = \sigma^2$, then $\overline{X} = \dfrac{\sum X_i}{n}$ is a random variable w/ $\mathbb{E}(\overline{X}) = \mu$ and $V(\overline{X}) = \dfrac{\sigma^2}{n}$

$\bar{X}$ more likely to be w/in larger sample size due to decreased variability.

NOTE: $(X_1 + X_2) - (Y_1 + Y_2 + Y_3) \neq 2X - 3Y$

Discrete Probability Distributions

X	Interpretation of X	Pmf: $P(X = x)$	$\mathbb{E}(x)$	$Var(x)$
Binomial	No of successes x in n trials	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$np(1-p)$
Geometric	No of trials x until 1st success	$(1-p)^{x-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negative binomial	No of trials x until rth success	$\binom{x-1}{r-1}(1-p)^{x-r}p^r$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
Hypergeometric	Pop size N, n draws, M success states, x successes, no replacement	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	$\frac{nM}{N}$	$\left(\frac{N-n}{N-1}\right)np(1-p)$
Poisson	No of arrivals x in fixed interval w/ expected value λ	$\frac{e^{-\lambda}\lambda^x}{x!}$	λ	λ

Properties of Discrete Probability Distributions

- Geometric dist'n has lack of memory property:
  $$P(X < t + \Delta t | X > t) = \frac{P(t < X < t + \Delta t)}{P(X > t)} = P(X < \Delta t)$$
- Geometric dist'n is special case of -ve binom dist'n w/ $r = 1$.
- Binomial $\approx$ hypergeometric if sampling small part of pop'n.
- Poisson $\approx$ binomial dist'n $\implies n$ is large and p is small.

	Binomial	-ve Binomial	Poisson
Trials	constant (n)	variable (X)	∞
Successes	variable (X)	constant (r)	variable (X)
P(success)	constant (p)	constant (p)	constant ($p = \lambda/n$)

Binomial	If a coin is tossed 20 times, what is the probability heads comes up exactly 14 times? $(x = 14,\ n = 20,\ p = 0.5)$
Geometric	If a coin is repeatedly tossed, what is the probability the first time heads appears occurs on the 8th toss? $(x = 8,\ p = 0.5)$
Negative binomial	If a coin is repeatedly tossed, what is the probability the third time heads appears occurs on the 9th toss? $(x = 9,\ r = 3,\ p = 0.5)$
Hypergeometric	5 pandas caught, tagged, and released into the pop. After mixing, a random sample of 10 animals selected. Suppose there are 25 animals in the region. What is the probability that exactly 2 of the caught animals are tagged? $(N = 25,\ n = 10,\ M = 5,\ x = 2)$
Poisson	Floods occur once every 100 years on avg in Canada. What is the probability there will be 4 floods in 100 years? $(\lambda = 1,\ x = 4)$

Continuous Probability Distributions

Distribution	Pdf: $f(x)$	$\mathbb{E}(X)$	$Var(X)$	Cdf: $F(x)$
Uniform	$\frac{1}{b-a}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\frac{x-a}{b-a}$
Exponential	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$1-e^{-\lambda x}$
Normal	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$\Phi(z)=P(Z\leq z)$

- Exponential dist'n has lack of memory property: $P(X<t+\Delta t|X>t)=P(X<\Delta t)$
- Exponential X is length of interval until next success in a Poisson process; $P(X)$ is probability of next success.
- Exponential λ is rate (i.e., per unit time); it's the SAME as that for Poisson.
- Poisson and exponential interchangeable if succcess hasn't happened yet ($N=0$).
- If $N\neq 0$, use Poisson (i.e., number of successes in interval): $\lambda=\lambda_x x$ where x is the interval.
- If success already happened, must use exponential.

	Poisson	Exponential
Length of interval	constant	variable (X, until 1st success)
No of successes	varianble (N)	constant (1)

$X\sim N(\mu,\sigma^2)$, $\quad Z\sim N(0,1)$ $\qquad Z=\frac{X-\mu}{\sigma},\quad z=\frac{x-\mu}{\sigma},\quad x=\mu+z\sigma$ $\qquad P(X\leq x)=P\left(\frac{X-\mu}{\sigma}\leq\frac{x-\mu}{\sigma}\right)=P(Z\leq z)$

Normal Approximation to Binomial

If $X\sim B(n,p)$, then $P(a\leq X\leq b)\approx P(a-0.5\leq Y\leq b+0.5)$ where $Y\sim N\left(\mu=np,\ \sigma^2=np(1-p)\right)$, $\quad np\geq 5$ and $n(1-p)\geq 5$

. . . .

Sampling Distributions

CLT: If n is sufficiently large, the sampling distribution of $\bar{x}$ is approximately normal, even if the pop'n dist'n is not itself normal.

Confidence lvl: Success rate of the method used to construct the interval: $100(1-\alpha)\%$. If $\alpha=0.05$, 95/100 samples will capture μ.

Lower confidence = narrower interval. Higher confidence = wider interval.

$\text{CI} = \text{estimator} \pm \underbrace{(\text{critical value})\times(\text{standard error})}_{\text{margin of error}}$ $\qquad \text{test statistic} = \frac{\text{estimate}-H_0}{\text{se(estimate)}}$

Decision		Actual situation: H_0 true	Actual situation: H_0 false
	Don't reject H_0	correct	type II, β
	Reject H_0	type I, α	correct

- α = significance = $P(H_0$ rejected $\mid H_0$ true)
- $1-\alpha$ = confidence
- $\beta=P(H_0)$ not rejected $\mid H_0$ false)
- $1-\beta$ = power of the test

$$\text{Evidence against } H_0\text{:}\begin{cases}0<p<0.01 & \text{convincing to strong}\\ 0.01<p<0.05 & \text{strong too moderate}\\ 0.05<p<0.1 & \text{moderate to suggestive but inconclusive}\\ 0.1<p<1 & \text{little to no}\end{cases}$$

Note: you can find sample size using $V(\hat{\theta})$, where $\hat{\theta}$ is a generic estimator for θ with $\mathbb{E}(\hat{\theta})=\mu_{\hat{\theta}}$ and SD $=\sigma_{\hat{\theta}}$.

Sample mean	Sample prop'n	Difference in sample means
$\mathbb{E}(\bar{X})=\mu$ $V(\bar{X})=\frac{\sigma^2}{n}$ Normal if $n\geq 30$ OR pop'n normal	$\mathbb{E}(\hat{p})=p$ $V(\hat{p})=\frac{p(1-p)}{n}$ Normal if $np\geq 5$ and $n(1-p)\geq 5$	$\mathbb{E}(\bar{X}_1-\bar{X}_2)=\mu_1-\mu_2$ $V(\bar{X}_1-\bar{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ Normal if $n_1\geq 30$ & $n_2\geq 30$ OR pop'n normal

Inferences About μ (σ known)

Asn's: $n \geq 30$ or pop'n is normal

Statistic: $z_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ CI: $\bar{x} \pm z_{\alpha/2}\left(\dfrac{\sigma}{\sqrt{n}}\right)$

Upper bound: $\mu \leq \bar{x} + z_\alpha\left(\dfrac{\sigma}{\sqrt{n}}\right)$

Lower bound: $\bar{x} - z_\alpha\left(\dfrac{\sigma}{\sqrt{n}}\right) \leq \mu$

Inferences About μ (σ unknown)

Asn's: random sample, $n \geq 30$ or pop'n normal

Statistic: $t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ CI: $\bar{x} \pm t_{\alpha/2,\ n-1}\left(\dfrac{s}{\sqrt{n}}\right)$

Upper bound: $\mu \leq \bar{x} + t_{\alpha,\ n-1}\left(\dfrac{s}{\sqrt{n}}\right)$

Lower bound: $\bar{x} - t_{\alpha,\ n-1}\left(\dfrac{s}{\sqrt{n}}\right) \leq \mu$

Inferences About μ_d for Paired Data

Asn's: paired and random samples, $n \geq 30$ or pop'n normal

Statistic: $t_0 = \dfrac{\bar{d} - \delta_0}{s_d/\sqrt{n}}$ CI: $\bar{d} \pm t_{\alpha/2,\ n-1}\left(\dfrac{s_d}{\sqrt{n}}\right)$

Inferences About $\mu_1 - \mu_2$ (indep, V known)

Statistic: $z_0 = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ CI: $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

Inferences About $\mu_1 - \mu_2$ (indep, V unknown, equal V)

Asn's: $n_1 \geq 30$ and $n_2 \geq 30$ or both pop'ns normal, $\dfrac{s_{\max}}{s_{\min}} < 2$

N.B., $n_1 \geq 15$ & $n_2 \geq 15$ to check equal variance

$\mathrm{df} = n_1 + n_2 - 2$

Statistic: $t_0 = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\mathrm{se}(\bar{x}_1 - \bar{x}_2)}$ CI: $(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\mathrm{df}} \times \mathrm{se}(\bar{x}_1 - \bar{x}_2)$

$s_p = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$ $\mathrm{se}(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Inferences About $\mu_1 - \mu_2$ (indep, V unknown, unequal V)

Asn's: $\dfrac{s_{\max}}{s_{\min}} > 2$ $\mathrm{df} = \min\{n_1, n_2\} - 1$

Statistic: $t_0 = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ CI: $(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\ \mathrm{df}}\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$

Inferences About p

Asn's: $np_0 \geq 5$ and $n(1 - p_0) \geq 5$

Statistic: $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$ CI: $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1 - \hat{p})}{n}}$

Upper bound: $p \leq \hat{p} + z_\alpha\sqrt{\dfrac{\hat{p}(1 - \hat{p})}{n}}$

Lower bound: $\hat{p} - z_\alpha\sqrt{\dfrac{\hat{p}(1 - \hat{p})}{n}} \leq p$

Inferences About $p_1 - p_2$

Asn's: $n_1\hat{p}_1$, $n_1(1 - \hat{p}_1)$, $n_2\hat{p}_2$, $n_2(1 - \hat{p}_2)$ all ≥ 5

Statistic: $z_0 = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$\hat{p} = \dfrac{\text{successes}}{\text{total}} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \dfrac{x_1 + x_2}{n_1 + n_2}$

CI: $(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$

Simple Linear Regression (SLR)

Model: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ $Y|x \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$

$\mathbb{E}(Y|x) = \mathbb{E}(\beta_0 + \beta_i x_i + \epsilon) = \beta_0 + \beta_1 x$ $\mathbb{E}(\epsilon) = 0$

$V(Y|x) = V(\beta_0 + \beta_1 x + \epsilon) = \sigma^2$ $V(\beta_0 + \beta_1 x) = 0$

dist'n of ϵ is normal $\sigma_\epsilon = \text{const for all } x$

Least Squares Fitting

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ Residuals: $\epsilon_i = y_i - \bar{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$

$\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}} = \dfrac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)} = r\dfrac{s_y}{s_x}$ $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$

- $S_{xx} = \sum(x_i - \bar{x})^2 = \sum x^2 - \dfrac{1}{n}\left(\sum x\right)^2$
- $S_{yy} = \sum(y_i - \bar{y})^2 = \sum y^2 - \dfrac{1}{n}\left(\sum y\right)^2$
- $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum xy - \dfrac{1}{n}\sum x \sum y$

ANOVA for SLR

$F_0 = \dfrac{\mathrm{SSR}/1}{\mathrm{SSE}/(n-2)} = \dfrac{\mathrm{MSR}}{\mathrm{MSE}}$ $(t_{n-2})^2 = F_{1,n-2}$

$$\overbrace{\sum_{i=1}^n (y_i - \bar{y})^2}^{SST} = \overbrace{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}^{SSR} + \overbrace{\sum_{i=1}^n (y_i - \hat{y}_i)^2}^{SSE}$$

$\mathrm{MSR} = \mathrm{SSR}/1$ $\mathrm{MSE} = \mathrm{SSE}/(n-2)$

	df	SS	MS	F	p-value
Regression	1	SSR	MSR	$\frac{MSR}{MSE}$	$P(F_{1,n-2} > F_0)$
Residual	$n-2$	SSE	MSE		
Total	$n-1$	SST			

Inferences in Regression

Statistic: $t_0 = \dfrac{\hat{\beta}_1 - b_1}{se(\hat{\beta}_1)} \sim t_{n-2}$ $se(\hat{\beta}_1) = \sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}$

CI (slope): $\hat{\beta}_1 \pm t_{\alpha/2,n-2} \times se(\hat{\beta}_1)$ $\hat{\beta}_1 \sim N(\beta_1, \sigma^2/S_{xx})$

Error/model variance: $\hat{\sigma}^2 = \dfrac{\mathrm{SSE}}{n-2}$

Inferences on Regression Line

CI: $(\hat{\beta}_0 + \hat{\beta}_1 x^*) \pm t_{\alpha/2,n-2} \times \hat{\sigma}\sqrt{\dfrac{1}{n} + \dfrac{(x^* - \bar{x})^2}{S_{xx}}}$

PI: $(\hat{\beta}_0 + \hat{\beta}_1 x^*) \pm t_{\alpha/2,n-2} \times \hat{\sigma}\sqrt{1 + \dfrac{1}{n} + \dfrac{(x^* - \bar{x})^2}{S_{xx}}}$

Correlation

Coefficient of determination: $R^2 = \dfrac{SSR}{SST} = 1 - \dfrac{SSE}{SST}$

Pearson's sample correlation: $r = \dfrac{1}{n-1}\sum z_x z_y = \dfrac{Sxy}{\sqrt{S_{xx}S_{yy}}}$

$\hat{\beta}_1 = r\dfrac{s_y}{s_x}$

$z_x = \dfrac{x_i - \bar{x}}{s_x}$ $\qquad$ $z_y = \dfrac{y_i - \bar{y}}{s_y}$

$R^2 = r^2$ for single-variable regression

One-Way ANOVA

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ $\qquad$ $SST = SSTr + SSE$

$$F_0 = \frac{\text{SSTr}/(k-1)}{\text{SSE}/(N-k)} = \frac{\text{MSTr}}{\text{MSE}} \sim F_{k-1,N-k}$$

$$F_0 = \frac{\text{MS btw groups}}{\text{MS w/in groups}}$$

Assumptions:

1. Samples from diff pop'ns are random and indep.
2. Pop'ns all normally distributed.
3. Pop'ns all have same sd.

$N = kn_i$ $\qquad$ k is n(groups); n_i is n(obs in each group)

	df	SS	MS	F	p-value
Tx	$k-1$	SSTr	MSTr	$\frac{MSTr}{MSE}$	$P(F_{1,n-2} > F_0)$
Error	$N-k$	SSE	MSE		
Total	$N-1$	SST			

Sample Size for Desired Margin of Error

One proportion: $n = \left(\dfrac{z^*}{ME}\right)^2 p^*(1-p^*)$

One mean: $n = \left(\dfrac{z^*}{ME}\right)^2 \hat{\sigma}^2$

Tips

On the TI-36X Pro, `{distn}cdf` computes $P(X \leq x)$

Use `quartile.inc()` and `tdist()` on Excel

Round DOWN in the t-table

1-sided test: use $\alpha = \frac{1}{2}(1 - \text{confidence}) \iff$ one-sided CI

2-sided test: use $\alpha = 1 - \text{confidence} \iff$ two-sided CI

Lower-tailed test $\iff$ upper bound

Upper-tailed test $\iff$ lower bound