# Final Review Part 2
## Chapters 3 & 4

### Section 3A

Null Hypothesis $H_0$
$=, \leq, \geq$, Not Related, No Correlation, No Association

Alternative Hypothesis $H_A$
$\neq, <, >$, Related, Correlation, Association

Claim: What someone thinks is true about population

$H_A: >$ greater than — Right Tail Test

$H_A: <$ less than — Left Tail Test

$H_A: \mu \neq \$50$
$H_A: \pi_1 \neq \pi_2$
$H_A: \beta_1 \neq 0$
} Two Tailed Tests — $T, Z$

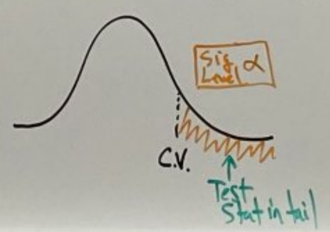
3 or more Populations
$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$
$H_0: \pi_1 = \pi_2 = \pi_3$
$H_A$: Categorical Variables Related
$H_A$: at least one $\neq$
} Right Tail — $\chi^2, F$

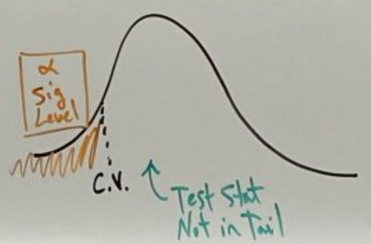
### Test Statistics Tail Rule

① Test Stat falls in tail determined by Critical value — Low P-value
- Sample Data Significantly disagrees with $H_0$
- $\bar{x}$ sig. disagrees with $\mu$
- $\hat{p}_1$ sig. disagrees with $\hat{p}_2$
- $r$ sig. different from 0
- observed counts sig disagree Expected counts
- Variance Between Sig Higher Variance within

② Test Statistic Does Not fall in a tail determined by Critical value — High P-value
- Sample data does not significantly disagree with $H_0$.
- $\bar{x}$ and $\mu$
- $\hat{p}_1$ and $\hat{p}_2$
- $r$ and 0
- observed counts Expected counts
- Variance Between Variance within

} Not Significantly Different

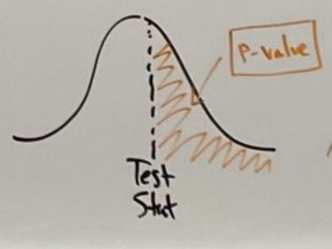


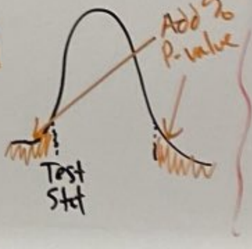




### P-value:

If $H_0$ is true, the Probability of getting the Sample Statistic or more extreme because of Sampling Variability (Random Chance)

* Convert P-value into %
* Compare to Significance Level $\alpha$ (1% 5% 10%)
* Low P-value: Lower than $\alpha$ — Reason Sample stat disagrees with $H_0$ unlikely to be sampling variability
* High P-value: Higher than $\alpha$ — Sampling Variability could be reason Sample Statistic disagrees with $H_0$



## Conclusions

Passing Conditions

Low P-value

"There (is) or (is not) Significant Evidence to (reject) or (support) the claim"

$H_0$ Claim

$H_A$ Claim

## Claim, Evidence

High P-value

| | | | | |
|---|---|---|---|---|
| 1 Pop. Mean<br>2 Pop. Mean | $H_0: \mu = 12$<br>$H_a: \mu_1 = \mu_2$ | $n \geq 30$ or Normal<br>Random, Indep. Observ. | T-test Stat (# standard errors)<br>$df = n-1$ | |
| 1 Pop. % Proportion<br>2 Pop. % Proportion | $H_0: \pi = .6$<br>$H_0: p_1 = p_2$ | At least 10 success<br>At least 10 failures<br>Random, Indep. Observ. | Z-test Stat (# Standard errors) | |
| Goodness of Fit (3 or more %) | $H_0: \pi_1 = \pi_2 = \pi_3 = \pi_4$<br>$H_0: p_1 = p_2 = p_3 = p_4$ | Expected Counts $\geq 5$<br>Random, Indep. Observ. | Chi Square test Stat $\chi^2$<br>$df = k-1$ | (observed counts vs Expected counts) |
| ANOVA<br>Categ./Quant Related? | $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ | Each group $n \geq 30$ or normal<br>Standard Deviations close<br>Random, Indep Observ. | F test stat<br>df between (numerator) $= k-1$<br>df within (denominator) $= n-k$ | (variance between Sig. higher than variance within?) |
| Correlation Regression (Quantitative Relationship Linear) | $H_0: \rho = 0$ $(\beta_1 = 0)$ | Random, Indiv. Indep.<br>$n \geq 30$<br>Scatterplot linear pattern<br>No outliers<br>Residual Plot evenly spread<br>Histogram of residuals Normal? Centred at zero | T-test Statistic | (# Standard errors between slope and zero) |

Categorical Association Test

$H_0$: Categ. not related

$H_A$: Categ are related

Expected Counts $\geq 5$

Random

Indiv. observ independent

Chi-Square $\chi$ test stat

$df = (r-1)(c-1)$

$\%$ could same in all 50 states? Goodness of Fit

Type of car (Categ) is related to # minutes to CAC (Quant)? ANOVA