Hypothesis Testing 101

Type of Sampling Distribution	One Mean	One Proportion	Difference between 2 independent sample means	Difference between 2 independent sample proportions
Center	$\mu_{\bar{x}} = \mu_x$	$\mu_{\hat{p}} = p$	$\mu_{\bar{x}} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2}$	$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$
Spread	$\sigma^2_{\bar{x}} = \frac{\sigma^2_x}{\sqrt{n^2}}$ $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$	$\sigma^2_{\hat{p}} = \frac{p(1-p)}{n}$ $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$	$\sigma^2_{\bar{x}} = \frac{\sigma^2_{\bar{x}_1}}{n_1} + \frac{\sigma^2_{\bar{x}_2}}{n_2}$ $\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}$	$\sigma^2_{\hat{p}_1 - \hat{p}_2} = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$ $\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
Shape	*IF n is sufficiently large, THEN sampling distribution of sample mean will be approximately normally distributed, no matter the shape of parent population distribution.*	*IF* $np \geq 10$ *AND* $n(1-p) \geq 10$, *THEN sampling distribution will be approximately normally distributed.*	*IF n is sufficiently large, THEN sampling distribution of difference between two independent sample means will be approximately normally distributed.*	*IF* $n_1 p_1 \geq 10, n_1(1-p_1) \geq 10, n_2 p_2 \geq 10,$ and $n_2(1-p_2) \geq 10$ *THEN sampling distribution of difference between two independent sample proportions will be approximately normally distributed.*
Assumptions & Conditions	1. *Random sample of observations* 2. *Independent observations* 3. $10n < N$ 4. *Must be roughly normal if* σ *is not known and/or n is small*	1. *Random sample of observations* 2. *Independent observations* 3. $10n < N$ 4. $np \geq 10$, $nq \geq 10$	1. *Random samples of observations from two independent populations* 2. *Independent observations* 3. $10n < N$	1. *Random samples of observations from two independent populations* 2. *Independent observations* 3. $10n < N$

Hypothesis Testing 101

<table>
<tr>
<td>Tests</td>
<td>z-test if σ is known & n > 40 ish

t-test if σ is not known and/or n is small</td>
<td>1-prop z-test

$z = \frac{obs - exp}{SE} = \frac{\hat{p} - p}{\sigma_{\hat{p}}}$</td>
<td>$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$
if σ's are known

$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
if σ's are not known</td>
<td>$z = \frac{(\hat{p}_1 - \hat{p}_2) - (\pi_1 - \pi_2)}{\sqrt{\hat{p}_c(1 - \hat{p}_c)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$\hat{p}_c = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$
$= \frac{combined\_successes}{combined\_sample\_size}$</td>
</tr>
<tr>
<td>Intervals</td>
<td>One Proportion z-Test →

$z = \frac{statistic - parameter}{SE}$

$z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}}$</td>
<td>One Proportion z-Interval →
sample statistic ± critical value * standard deviation

$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$</td>
<td>$(\bar{x}_1 - \bar{x}_2) \pm t^*_{df} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$</td>
<td>$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_2)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$</td>
</tr>
<tr>
<td colspan="3">If we don't know

μ - we estimate it with $\bar{X}$

σ – we estimate it with S_X

p(π) – we estimate it with $\hat{p}$

if we use these estimates to calculate the variability for a sampling distribution, we now call the <u>standard error</u>.</td>
<td colspan="2">So... If SD: $(\hat{p}) = \sqrt{\frac{p(1 - p)}{n}}$

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

And...If SD $(\bar{X}) = \frac{\sigma_X}{\sqrt{n}}$

Then the SE $(\bar{X}) = \frac{s_x}{\sqrt{n}}$</td>
</tr>
</table>

Hypothesis Testing 101

H: Hypotheses

State in symbols and in words

C: Conditions

Independent
Random
Large enough
Success/Failure

Be sure to state if you are able to use the Normal model or not!

T: Test statistic

Write the entire formula with correct symbols, including degrees of freedom (df) or name that test!
Evaluate the test statistic by writing in the values and having the calculator produce the numbers (including possibly, df's)
State "by calculator" in your answer

A: Alpha

Compare p-level to alpha, include a properly labeled sketch.

C: Conclude:

Cite the comparison of p-level to alpha AND state conclusion in context.

I: Introduce

A full sentence identifying the parameter in context and in symbol:

- *"I am creating a 99% confidence interval for μ, the mean radon level in ppm in houses in anytown."*

C: Conditions

List and check conditions as needed, including random sample, $n \leq .1N$, and evidence of normality if needed

- *(np & nq are at least ten, etc., or boxplot checked for symmetry, or n large, whatever your text asks)*

F: Formula

Write the entire formula with correct symbols, including df

C: Calculations

Write in the values, including the z- or t - critical value. Have the calculator produce the interval. *State "by calculator" in your answer*

I: Interpret.

Two sentences:

- One for the numbers in context
    - *I am 99% confident that the true mean radon levels in ppm lies in the interval between ??? and ???*
- One for the method
    - *99% of similiarly constructed intervals will contain the true mean.*