Pre Midterm

$Z = \frac{X-\mu}{\sigma}$	$\Phi(z) = P(Z \le z)$	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2\right)$

Central Limit Theorem (σ known)

$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$\overline{X}$ is sample mean, from a population with mean μ and variance σ^2.
Exception: if σ unknown but $n > 30$, then we assume $\sigma \approx s$.	

Student t (σ not known)

$t = \frac{\overline{x}-\mu}{s/\sqrt{n}}$	μ is population mean, $\overline{x}$ is sample mean, s is sample st. deviation, and n is sample size.
$t_{a,k}$: α is probability in range $(t_{a,k} + \infty)$, k is deg of freedom.	

Hypothesis Testing, C.I.

Procedure: 1) **Param. of interest** μ. 2) H_0 and H_1. 3) **Test Stats.** model: CLT. 4) Find critical region. 5) Find p value (min α that rejects H_0).

	2-Sided	Critical Region **RHS**	Critical Region **LHS**
H_1, δ	$H_1: \mu \ne \mu_0, \delta \ne 0$, $\boldsymbol{\mu_0}$ = false mean; $\boldsymbol{\mu} = \mu_0 + \delta$ = true mean	$H_1: \mu > \mu_0, \delta > 0$	$H_1: \mu < \mu_0, \delta < 0$
β	$P\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma} < Z < z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \approx \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$	$P\left(Z < z_\alpha - \frac{\delta\sqrt{n}}{\sigma}\right)$	$P\left(Z > -z_\alpha - \frac{\delta\sqrt{n}}{\sigma}\right)$

Power of Test $= 1 - \beta = P(\text{Correctly Reject False } H_0)$
Reduce α increases β. Increase $n \to \alpha, \beta$ *both* decrease.
For **one-sided hypothesis**, $H_0: \mu = \mu_0$, assuming the irrelevant side is included, **and we wish to reject** $\boldsymbol{H_0}$.

Find Sample Size: Let $\beta = \Phi(-z_\beta)$, $-z_\beta \approx z_{\alpha/2} - \delta\sqrt{n}/\sigma$

Two Sided Test	One Sided Test
$n \approx (z_{\alpha/2} + z_\beta)^2\sigma^2/\delta^2$	$n \approx (z_\alpha + z_\beta)^2\sigma^2/\delta^2$

Confidence Interval (CI): CI contains true param.
Confidence Level: probability that μ is within CI

σ unknown: $P\left(\overline{x} - t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}\right) = 1 - \alpha$

σ known	2-sided $P(L \le \mu \le U)$	**L-Conf. Bound** $P(\mu \ge L)$	**U-Conf. Bound** $P(\mu \le U)$
Formula	$P\left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$	$P(\mu > \overline{x} - z_\alpha\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$	$P(\mu < \overline{x} + z_\alpha\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$

Decisions for Two Samples: $H_0: \mu_1 - \mu_2 = 0$, $H_1: \mu_1 - \mu_2 \ne 0$

	σ_1, σ_2 known	$\sigma_1 = \sigma_2$ unknown	$\sigma_1 \ne \sigma_2$ unknown
Test Statistics	$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}}$ $S_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$ $k = n_1 + n_2 - 2$	$\boldsymbol{T^*} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$ $k = \text{round}\left[\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}\right]$
C.I. (replace $\alpha/2$ with α for 1 sided bound)	$(\overline{X_1} - \overline{X_2}) \pm z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$(\overline{X_1} - \overline{X_2}) \pm t_{\alpha/2,k} \cdot S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$(\overline{X_1} - \overline{X_2}) \pm t_{\alpha/2,k_{nasty}} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Factorial Experiment

	Formula	Explanation
2^2	$A = (1/2n)[a + ab - b - (1)]$, $AB = (1/2n)[ab + (1) - a - b]$	$a, ab, b(1)$ are **total system responses**, n is # trials.
2^3	Refer to the contrast table for signs of $a, b, c, ab, ac, bc, abc, (1)$ when finding main and interaction effects.	

Main Effect $\boldsymbol{A} = (A_{\text{high}})_{\text{avg}} - (A_{\text{low}})_{\text{avg}}$ **Label** in contrast table (e.g. *a, b, ab*) indicates a *high* treatment combination.
Interaction Effect $\boldsymbol{AB}$ = Average Response(A, B same level) − Average response(A, B different levels).

Hypo. test is used to decide whether **t ratio** values are significant or not. $\begin{cases} H_0: \text{Effect} = 0 \\ H_1: \text{Effect} \ne 0 \end{cases}$ $\boldsymbol{t_{\text{ratio}}} = \frac{\text{Effect}}{s_e(\text{Effects})} \leftarrow$ normalize. Use t dist., deg. of freedom is $k = 2^\kappa(n-1)$

$s_e(\textbf{Effects}) = \sqrt{\frac{\hat{\boldsymbol{\sigma}}^2}{n \cdot 2^{\kappa-2}}}$	$\hat{\boldsymbol{\sigma}}^2 = \frac{1}{2^\kappa}\sum_{i=1}^{2^\kappa}\hat{\boldsymbol{\sigma}}_i^2$	$\hat{\boldsymbol{\sigma}}_i^2 = \frac{\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2}{(n-1)}$	where κ is #factors, n is #trials; $\hat{\sigma}^2$ is avg. of var. at each treatment: $\hat{\sigma}_i^2$ is the var. at each treatment.

Effect $\pm 2 \times s_e(\text{Effects}) \approxeq 95\%$ **C. I.** on Effect **If 95% CI Contains zero, effect is *not* significant.**

Method of Steepest Accent: 1st-order method to move to *vicinity* of optimum.
1st-Order Response Surface: 1st-order, lin. reg. model of response surface

$$Y = \overline{y} + \left(\frac{A}{2}\right)x_1 + \left(\frac{B}{2}\right)x_2 + \left(\frac{AB}{2}\right)x_1x_2$$
$$\hat{Y} = \overline{y} + \left(\frac{A}{2}\right)x_1 + \left(\frac{B}{2}\right)x_2$$

From the center of the surface, on the x_1x_2 plane, notice that the ratio B/A is the slope of the line describing the path to a more optimal response.

Linear Regression

$\hat{y}_i = a + b\,x_i$	$\hat{y}_i$ = Predicted Value y_i = Observed Value $\bar{y}$ = Average of observed e_i = Error SS_E = Err. Sum of Sqrs SS_R = Regression Sum of Sqrs. SS_T = Total Sum of Squares S_e = Standard Err.	**Best Fit Line**: $b = S_{xy}/S_{xx}$	$a = \bar{y} - b\bar{x}$
$e_i = y_i - \hat{y}_i$			
$SS_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$			
$SS_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$			

True Regression Line is defined as $y = \alpha + \beta x$

Judging Fit of Data: $H_0: \beta = 0, H_1: \beta \neq 0$ **Goal**: reject H_0

Std Err Estimator Approach: model as t distribution $T = \dfrac{b - \beta}{S_e/\sqrt{S_{xx}}}$

Reject H_0 if p value < given α (evidence of lin. relationship)

$$SS_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 \approx \sum(y_i^2) - \frac{(\sum y_i)^2}{n} = SS_R + SS_E$$

$$S_e \approx \sqrt{\hat{\sigma}^2} \quad \text{where} \quad \hat{\sigma}^2 = \frac{SS_E}{n-2}$$

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \approx \sum(xy) - \frac{\sum(x)\sum(y)}{n}$$

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 \approx \sum(x^2) - \frac{\sum(x)\sum(x)}{n}$$

C.I.: *population mean of* y *at* x_0, denoted $\mu(y|_{x_0}) = \alpha + \beta x_0$

$$P\left(L < \mu(y|_{x_0}) < U\right) = 1 - \alpha$$

$$L, U = a + bx_0 \pm t_{\alpha/2, n-2} \cdot S_e \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

Prediction Interval: predicts Y_0 at x_0 observation.

$$Y_0|_{x_0} = a + bx_0 \pm t_{\alpha/2, n-2} \cdot S_e \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

Formula diff. from C.I. due to err. From new measurement.

Coeff. of Determination is the proportion of SS_T explained by use of regression model. $R^2 = \dfrac{SS_R}{SS_T} = 1 - \dfrac{SS_E}{SS_T}$ When R^2 is close to 1, it's a good fit of data.

Uncertainty Analysis

Uncertainty w_R of result $R = f(x_1, x_2, \dots, x_n)$, with $x_1 \pm w_1, \dots, x_n \pm w_n$

$$(w_R)_{max} = \left|w_1 \frac{\partial R}{\partial x_1}\right| + \cdots + \left|w_n \frac{\partial R}{\partial x_n}\right| \qquad w_R = \left(\sum_{i=1}^{n}\left[w_i \cdot \frac{\partial R}{\partial x_i}\right]^2\right)^{1/2}$$

Special case: $R = C x_1^a x_2^b \dots x_n^N$

$$\frac{w_R}{R} = \sqrt{\left(\frac{w_1 a}{x_1}\right)^2 + \left(\frac{w_2 b}{x_2}\right)^2 + \cdots + \left(\frac{w_n N}{x_n}\right)^2}$$

$w_{\bar{x}} = \left(B_x^2 + P_{\bar{x}}^2\right)^{1/2}$ **Systematic Err. B_x** Usually given as **accuracy full scale**.

Random Err. $P_x = \pm t \cdot s_x$ [single pt.] $P_{\bar{x}} = \pm\, t \cdot \dfrac{s_x}{\sqrt{n}}$ t is the Student t value for 95% conf. level ($\alpha = 0.025$), s_x is S.D. (known), n is the sample size (# prev trials)

Cali. Curve: true vs measured. Best fit line is **cali. eq.** **Deviation** = True − Cali(Measured)

Deviation Plot: deviation vs true val. **Max Uncertainties**: lowest and highest deviation

Joint Probability

r denotes system success probability, $r_1, \dots, r_k$ are success probability of independent sub components

$$P(A \vee B) = P(A) + P(B) - P(A) \cdot P(B)$$

Series System	Parallel System
$r = \prod_{i=1}^{k} r_i$	$r = 1 - \prod_{i=1}^{k}(1 - r_i)$

$X_1, X_2, \dots, X_n$	Function Y	$E(Y)$	$V(Y)$
Independent	$c_0 + c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$	$c_0 + c_1\mu_1 + c_2\mu_2 + \cdots + c_n\mu_n$	$c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \cdots + c_n^2\sigma_n^2$
Dependent	$Y = c_0 + c_1 X_1 + c_2 X_2$	$c_0 + c_1\mu_1 + c_2\mu_2$	$c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2c_1c_2\,\mathbf{Cov}(X_1, X_2)$

Covariance Cov$(X_1, X_2) = [E(X_1X_2) - \mu_1\mu_2]$, when X_1, X_2 are independent, $\mathrm{Cov}(X_1, X_2) = 0$.

$$\textbf{correlation } \rho_{X_1, X_2} = \frac{E(X_1X_2) - \mu_1\mu_2}{\sqrt{\sigma_2^2\sigma_2^2}} = \frac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\sigma_1^2\sigma_2^2}} \in [-1, +1]$$