Definition. ***Composite Hypotheses*** *When alternative hypothesis is of form $\mathcal{H}_1 : \theta \in \Theta_1$, where Θ_1, the composite alternative, consists of more than a single possible value.*

Definition. ***Power function*** *of a statistical test os*

$$\pi(\theta^*) = \mathbb{P}\,(reject\ \ \mathcal{H}_0 \,|\, \theta = \theta^*)$$

where $\theta^* \in \Theta_1$.

Example. Likelihood ratio test (Right tail test) for $\mathcal{H}_0 : \mu = \mu_0$ vs $\mathcal{H}_1 : \mu > \mu_0$ for Normal data with known variance,

$$\pi(\mu^*) = \mathbb{P}(\underline{X} \in \mathcal{C} | \mu = \mu^*) = 1 - \Phi\left(\frac{-\sqrt{n}(\mu^* - \mu_0)}{\sigma} + z_{1-\alpha}\right)$$

where $\mu^* > \mu_0$

1. larger n, test becomes more powerful
2. The further apart null hypothesis and alternatives are, the greater the power
3. larger σ, less powerful the test
4. smaller α, the smaller the power (the α-β trade-off)

Now we can maximize power π by increasing sample size n. We keep the probability of Type II error at less than β (i.e. having power greater than $1-\beta$) beyond differences larger than $\delta = \mu_1 - \mu_0$

$$\begin{aligned}
1 - \beta \leq \pi(\mu_1) &= 1 - \Phi(-\frac{\sqrt{n}(\mu_1 - \mu_0)}{\sigma} + z_{1-\alpha}) \\
\Rightarrow -\frac{\sqrt{n}(\mu_1 - \mu_0)}{\sigma} &= z_{1-\alpha} \leq z_\beta = -z_{1-\beta} \\
\Rightarrow n \geq \left\{\frac{\sigma(z_{1-\alpha} + z_{1-\beta})}{\mu_1 - \mu_0}\right\}^2
\end{aligned}$$

Definition. ***Uniformly Most Poweful (UMP) Test*** *A test that is MP for every simple alternative $\theta \in \Theta_1$ is UMP. Consider testing $\mathcal{H}_0 : \theta = \theta_0$ vs. $\mathcal{H}_1 : \theta \in \Theta_1$ (a composite alternative). We say that a test at level α with power function $\pi(\theta)$ is a uniformly most powerful (UMP) test, if for any other test at level α with power function $\pi'(\theta)$, we have $\pi'(\theta) \leq \pi(\theta)$ for all $\theta \in \Theta_1$.*

Example. For $X_1, \cdots, X_n \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma^2)$, the rejection region

$$\mathcal{C} = \left\{\underline{x} \in \mathbb{R}^n : \overline{x} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right\}$$

for testing simple hypothesis $\mathcal{H}_0 : \mu = \mu_0$ vs. $\mathcal{H}_1 : \mu = \mu_1\,(\mu_1 > \mu_0)$ is the most powerful test at level α by Pearson-Neyman Lemma. Since rejection region does not depend on μ_1, it is therefore the most powerful test for any $\mu \in \Theta_1$ where $\Theta_1 = (\mu_0, \infty]$. Hence the likelihood ratio test with above rejection region is the UMP test for testing $\mathcal{H}_0 : \mu = \mu_0$ vs. one-tailed alternative $\mathcal{H}_1 : \mu > \mu_1$.
Consider testing $\mathcal{H}_0 : \mu = \mu_0$ vs. $\mathcal{H}_1 : \mu \neq \mu_0$. Because the respective one tailed test is UMP, the test for two-sided alternatvie is not same for every alternative, hence it is not UMP. Usually, composite hypothesis has no UMP test.

Definition. ***p-value*** *the probability of observing an effect at least as extreme as the one in observed data, assuming the truth of $\mathcal{H}_0$.*

$$\textit{p-value} = \mathbb{P}(\textit{Type I Error}) = \mathbb{P}(\underline{X} \in \mathcal{C} \,|\, \theta = \theta_0) \qquad \textit{where} \qquad \mathcal{C} = \{T(\underline{X}) \geq t(\underline{x})\}$$

where $T(X)$ the test statistic. Hence if p-value is very low, then $\mathcal{H}_0$ is most likely false. Alternatively, it is the minimum α for which $\mathcal{H}_0$ will be rejected.

$$\textit{Reject } \mathcal{H}_0 \textit{ at level } \alpha \iff p - value \leq \alpha$$

Remark. This avoids having to re-calculate the rejection region based on different α. Now we reject $\mathcal{H}_0$ at α if and only if p-value is less than α. However, this does not prove that the $\mathcal{H}_1$ is true. Also, The p-value does not support reasoning about the probabilities of hypotheses but is only a tool for deciding whether to reject the null hypothesis.

Definition. ***Composite Null Hypothesis*** *It is more sensible to write*

$$\begin{cases} \mathcal{H}_0 : \mu \leq \mu_0 \\ \mathcal{H}_1 : \mu > \mu_0 \end{cases} \qquad \begin{cases} \mathcal{H}_0 : \mu \geq \mu_0 \\ \mathcal{H}_1 : \mu < \mu_0 \end{cases}$$

for right-tailed and left-tailed test respectively. Correspondingly, we re-define alpha to be

$$\alpha = \sup_{\theta \in \Theta_0} \pi(\theta)$$

as an example, for right tailed test for $\mathcal{H}_0 : \mu \leq \mu_0$ vs $\mathcal{H}_1 : \mu > \mu_0$ in case of Normal distribution with known variance, we have

$$\pi(\mu^*) = 1 - \Phi\left(\frac{\sqrt{n}(\mu_0 - \mu^*)}{\sigma} + z_{1-\alpha}\right)$$

where $\mu^* \leq \mu_0$, Note $\pi(\mu^*)$ is monotonically increasing and achieves max at right endpoint when $\mu^* = \mu_0$, hence $\alpha = \sup_{\theta \in \Theta_0} \pi(\theta)$ holds.

Example. Two-tailed test for Normal Mean

1. **Rejection region** Large values of $|\overline{X} - \mu_0|$ is a strong evidence against $\mathcal{H}_0$

$$\mathcal{C} = \left\{|\overline{X} - \mu_0| \geq \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right\} = \left\{\overline{X} \leq \mu_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right\} \cup \left\{\overline{X} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right\}$$

Note we have $1 - \alpha/2$ instead of $1 - \alpha$ for one tailed test. In general, its easier to reject one tailed test than to reject two tailed test

2. **p-value**

$$\text{p-value} = \mathbb{P}\left(|\overline{X} - \mu_0| \geq |\overline{x} - \mu_0| \middle| \mu = \mu_0\right) = 2(1 - \Phi(\frac{|\overline{x} - \mu_0|}{\sigma/\sqrt{n}}))$$

We then plug in observed $\overline{x}$ and given parameter μ_0, σ, n to calculate the p-value. The result exactly doubles that of the one tailed test. Since α has to be larger than p-value for rejection to happen, it becomes harder to reject two tailed test than one tailed test (ex. an arbitrarily small p-value can be rejected by any α specified).

3. **Power Function**

$$\begin{aligned}
\pi(\mu^*) &= \mathbb{P}(\underline{X} \in \mathcal{C}|\mu = \mu^*) \\
&= \mathbb{P}(\overline{X} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}|\mu = \mu^*) + \mathbb{P}(\overline{X} \leq \mu_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}|\mu = \mu^*) \\
&= 1 - \Phi(-\frac{\sqrt{n}(\mu^* - \mu_0)}{\sigma} + z_{1-\alpha/2}) + \Phi(-\frac{\sqrt{n}(\mu^* - \mu_0)}{\sigma} - z_{1-\alpha/2})
\end{aligned}$$

A plot of $\pi(\mu^*)$ to μ^* resembles an upside down bell curve, where global minimum happens at $\mu^* = \mu_0$. Also note that two tail tests are not UMP test because one tail test are UMP test in their respective domains by Neyman-Pearson lemma.

4. **Confidence Intervals** Acceptance region and $(1 - \alpha)\%$ confidence interval for μ are equivalent.

$$\begin{aligned}
\text{Do not reject } \mathcal{H}_0 : \mu = \mu_0 &\iff \overline{X} \not\subseteq \mathcal{C} \\
&\iff \overline{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \leq \mu_0 \leq \overline{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \\
&\iff \overline{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \leq \mu_0 \leq \overline{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2} \\
&\iff \mu_0 \text{ contained in the } (1 - \alpha)\% \text{ confidence interval for } \mu
\end{aligned}$$

Confidence interval consists of precisely of all those values of μ_0 for which the null hypothesis $\mathcal{H}_0 : \mu = \mu_0$ is accepted, i.e. a set of plausible value for μ. If we reject $\mathcal{H}_0 : \mu = \mu_0$ then μ_0 is not a plausible value. In general, the set of all θ_0 for which $\mathcal{H}_0 : \theta = \theta_0$ would not get rejected in a two-tailed set at level α forms a $(1 - \alpha)\%$ *confidence set* for θ. Every confidence set has a corresponding two-tailed test.

Theorem. *Suppose that for every $\theta_0 \in \Theta$ there is a test at level α of the hypothesis $\mathcal{H}_0 : \theta = \theta_0$. Denote the acceptance region of the test by $A(\theta_0) = \Omega \setminus \mathcal{C}$, then the set*

$$I(\underline{X}) = \{\theta : \underline{X} \in A(\theta)\}$$

is a $100(1-\alpha)\%$ confidence region for θ. That is for every θ_0

$$\mathbb{P}(\theta_0 \in I(\underline{X}) \,|\, \theta = \theta_0) = 1 - \alpha$$

Remark. A confidece region for θ consists of all those value θ_0 for which the hypothesis that θ euqals θ_0 will be rejected at level α. The hypothesis that $\theta = \theta_0$ is accepted if θ_0 lies in the confidence interval I.