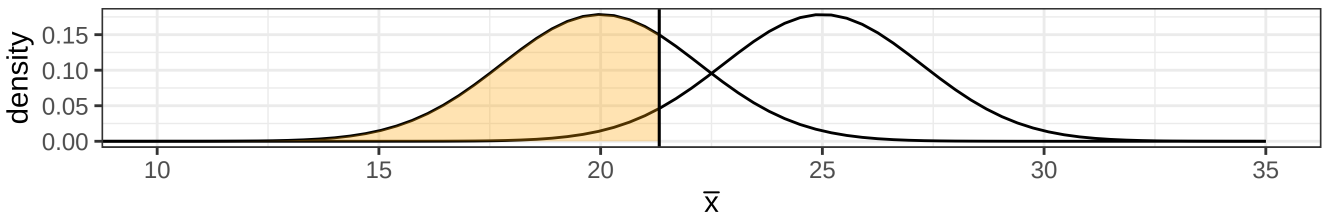
# Warm Up: Power Functions for Hypothesis Tests

- Data Model: $X_1, \dots, X_5 \overset{\text{i.i.d.}}{\sim} \text{Normal}(\theta, 5^2)$. We saw that the likelihood ratio test is *equivalent* to a test based on $\bar{x}$.

**1. Consider a test of the hypotheses $H_0 : \theta = 25$ vs. $H_A : \theta = 20$. Below is a picture showing the pdf $f_{\bar{X}|\theta}(\bar{x}|20)$ of a Normal$(20, 5^2/5)$ distribution and the pdf $f_{\bar{X}|\theta}(\bar{x}|25)$ of a Normal$(25, 5^2/5)$ distribution, along with a vertical line at $q_5^{null}$, the 5th percentile of the Normal$(25, 5^2/5)$ distribution.**

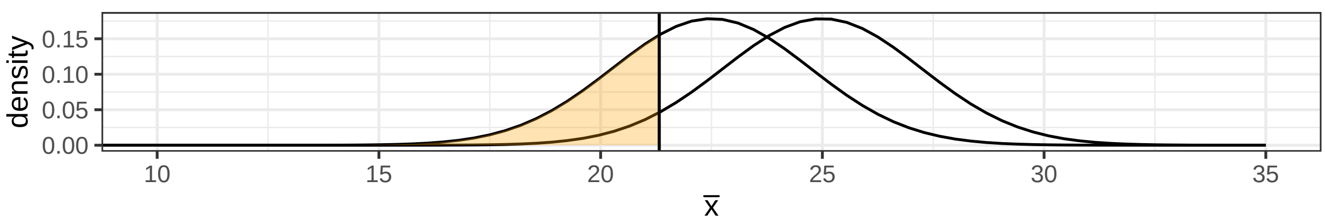
- **Shade in the area corresponding to $1-\beta$, the power of the likelihood ratio test if $H_A$ is correct.**
- **Show how you would calculate the power of the test as an integral of either $f_{\bar{X}|\theta}(\bar{x}|20)$ or $f_{\bar{X}|\theta}(\bar{x}|25)$**




$1 - \beta = \int_{-\infty}^{q_5^{null}} f_{\bar{X}|\theta}(\bar{x}|20) d\bar{x}$

**2. Suppose that instead we were testing the hypotheses $H_0 : \theta = 25$ vs. $H_A : \theta = 22.5$. Below is a picture showing the pdf $f_{\bar{X}|\theta}(\bar{x}|22.5)$ of a Normal$(22.5, 5^2/5)$ distribution and the pdf $f_{\bar{X}|\theta}(\bar{x}|25)$ of a Normal$(25, 5^2/5)$ distribution, along with a vertical line at $q_5^{null}$, the 5th percentile of the Normal$(25, 5^2/5)$ distribution.**

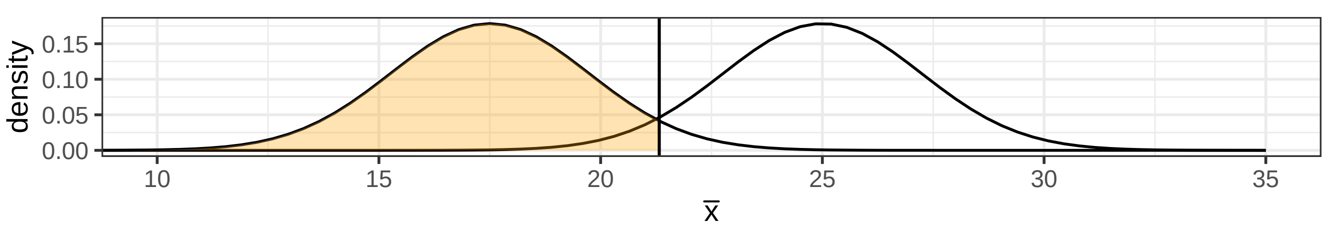
- **Shade in the area corresponding to $1-\beta$, the power of the likelihood ratio test if $H_A$ is correct.**
- **Show how you would calculate the power of the test as an integral of either $f_{\bar{X}|\theta}(\bar{x}|22.5)$ or $f_{\bar{X}|\theta}(\bar{x}|25)$**




$1 - \beta = \int_{-\infty}^{q_5^{null}} f_{\bar{X}|\theta}(\bar{x}|22.5) d\bar{x}$

**3. Suppose that instead we were testing the hypotheses $H_0 : \theta = 25$ vs. $H_A : \theta = 17.5$. Below is a picture showing the pdf $f_{\bar{X}|\theta}(\bar{x}|17.5)$ of a Normal$(17.5, 5^2/5)$ distribution and the pdf $f_{\bar{X}|\theta}(\bar{x}|25)$ of a Normal$(25, 5^2/5)$ distribution, along with a vertical line at $q_5^{null}$, the 5th percentile of the Normal$(25, 5^2/5)$ distribution.**

- **Shade in the area corresponding to $1-\beta$, the power of the likelihood ratio test if $H_A$ is correct.**
- **Show how you would calculate the power of the test as an integral of either $f_{\bar{X}|\theta}(\bar{x}|17.5)$ or $f_{\bar{X}|\theta}(\bar{x}|25)$**




$1 - \beta = \int_{-\infty}^{q_5^{null}} f_{\bar{X}|\theta}(\bar{x}|17.5) d\bar{x}$

**4. For which of the alternative hypotheses above ($\theta = 17.5$, $\theta = 20$, or $\theta = 22.5$) is the power of the test largest? For which is the power smallest?**

The power is largest for $\theta = 17.5$ and smallest for $\theta = 22.5$.

**Definition: power function**

The *power function* $K(\theta)$ for a test is the power of the test at $\theta$:

$K(\theta) = \int_R f_{\mathbf{X}|\theta}(\mathbf{x}|\theta) d\mathbf{x}$

where $R$ denotes the rejection region of the test (i.e., $R$ is the set of $\mathbf{x}$ such that the p-value is less than $\alpha$)

**Definition: Uniformly most powerful test**

A test with power function $K(\theta)$ if for any other test with power function $K'(\theta)$, $K(\theta) \geq K'(\theta)$ for all $\theta \in \Omega_0^c$ (i.e., it has at least as large of power as any other test for every possible parameter consistent with the alternative hypothesis).

**(Non-) Example**

Here is a plot of the power functions of 3 different tests of

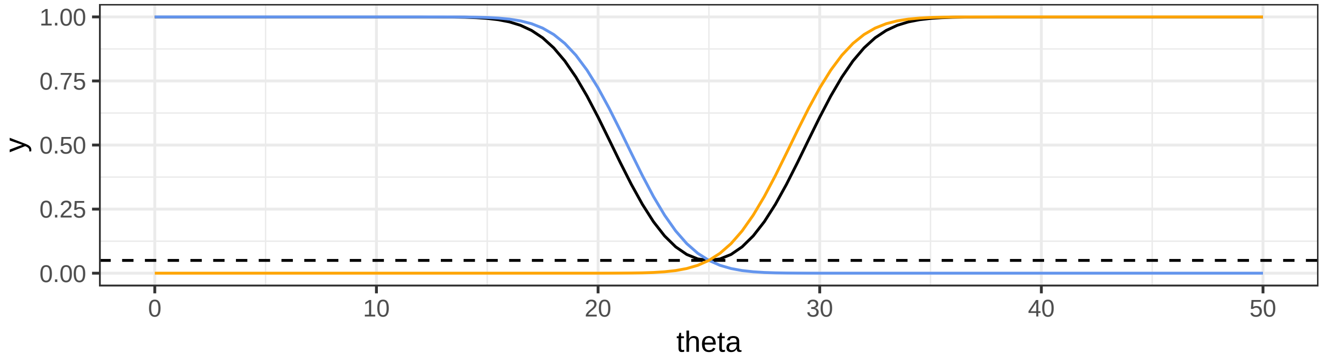
$H_0 : \theta = 25$ vs. $H_A : \theta \neq 25$

for our running example with $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \text{Normal}(\theta, 5^2)$.

- Shown in black is the power function for a test that rejects $H_0$ if $\bar{x} \leq q^{null}_{0.025}$ or $q^{null}_{0.975} \leq \bar{x}$
- Shown in blue is the power function for a test that rejects $H_0$ if $\bar{x} \leq q^{null}_{0.05}$
- Shown in orange is the power function for a test that rejects $H_0$ if $q^{null}_{0.95} \leq \bar{x}$

```r
calc_power_two_sided <- function(theta) {
  lower_cutoff <- qnorm(0.025, mean = 25, sd = sqrt(5))
  upper_cutoff <- qnorm(0.975, mean = 25, sd = sqrt(5))

  power <- rep(NA, length(theta))
  for(i in seq_along(theta)) {
    power[i] <- pnorm(lower_cutoff, mean = theta[i], sd = sqrt(5)) +
      pnorm(upper_cutoff, mean = theta[i], sd = sqrt(5), lower.tail = FALSE)
  }
  return(power)
}

calc_power_left_side <- function(theta) {
  lower_cutoff <- qnorm(0.05, mean = 25, sd = sqrt(5))

  power <- rep(NA, length(theta))
  for(i in seq_along(theta)) {
    power[i] <- pnorm(lower_cutoff, mean = theta[i], sd = sqrt(5))
  }
  return(power)
}

calc_power_right_side <- function(theta) {
  upper_cutoff <- qnorm(0.95, mean = 25, sd = sqrt(5))

  power <- rep(NA, length(theta))
  for(i in seq_along(theta)) {
    power[i] <- pnorm(upper_cutoff, mean = theta[i], sd = sqrt(5), lower.tail = FALSE)
  }

  return(power)
}

ggplot(data = data.frame(theta = c(0, 50)), mapping = aes(x = theta)) +
  stat_function(fun = calc_power_two_sided) +
  stat_function(fun = calc_power_left_side, color = "cornflowerblue") +
  stat_function(fun = calc_power_right_side, color = "orange") +
```

```
  ylim(0, 1) +
  geom_hline(yintercept = 0.05, linetype = 2) +
  theme_bw()
```




- By the Neyman-Pearson lemma, the left-sided test (blue) is most powerful for any $\theta_1 < \theta_0$
- By the Neyman-Pearson lemma, the right-sided test (orange) is most powerful for any $\theta_1 > \theta_0$
- Therefore, there is no test that is uniformly most powerful for a two-sided test about the mean of a normal distribution.