We are sampling from a population of fish, or more properly, of fish lengths. (This example is from Mayo (1985), but you do not need that paper for this.) A fish’s length, in inches, may be represented by variable X; that is, to each fish a value of X (like a little badge) is attached.

A random sample of size n is taken, $X = (X_1, \ldots, X_n)$, where each $X_i$ is distributed Normally with unknown mean $\mu$ and known standard deviation $\sigma = 2$. Suppose we want to test the hypotheses:

$$H: \mu = \mu_0 = 12, \text{ against } J: \mu > \mu_0.$$ 

$H$ is the “null” or “test” hypothesis, and $J$ the (composite) alternative hypothesis. The N-P test is a rule that tells us for each possible outcome $x = (x_1, \ldots, x_n)$ whether to reject or accept $H_0$. The rule is defined in terms of a test statistic or distance measure:

$$d(X) = \frac{X - \mu}{\sigma_x},$$

where $\bar{X}$ is the sample mean whose sampling distribution is also Normal with mean $\mu$ and standard deviation $\sigma_x = \sigma / \sqrt{n}$.

In order to test these claims, we observe, randomly, n fish (fish lengths), and average their lengths to get $\bar{X}$. We perform a one-sided test.

**Under the null hypothesis $H$:** $\bar{X}$ is Normal $(\mu, \sigma_x)$, and $\mu = 12$.

**Under the alternative hypothesis $J$:** $\bar{X}$ is Normal $(\mu, \sigma_x)$, and $\mu > 12$.

$\bar{X}_{obs}$ is the sample (observed) mean.

1. To calculate the **statistical significance level** of $\bar{X}_{obs}$
   a. Calculate $D_{obs} = \bar{X}_{obs} - \bar{X}_{obs}(\text{expected assuming } H) = \bar{X}_{obs} - 12$
   
   b. Put $D_{obs}$ in standard deviation units: $D_{obs} / \sigma_x$
   
   c. Find the area between 0 and $D_{obs} / \sigma_x$ on the Standard Normal distribution chart.
   
   d. .5 minus the value you get in c is the **statistical significance level** of the observed difference.

2. The **severity** with which $J$ passes this test with outcome $\bar{X}_{obs}$ equals 1 minus the statistical significance level of $D_{obs}$. See #6. below.
Problem Set #1: Find the statistical significance level of each of the following outcomes with the indicated sample size n:

a. n = 25:
   (i) $\bar{X}_{obs} = 12.6$
   (ii) $\bar{X}_{obs} = 12.8$
   (iii) $\bar{X}_{obs} = 13.$

b. n = 100:
   (i) $\bar{X}_{obs} = 12.2$
   (ii) $\bar{X}_{obs} = 12.4$
   (iii) $\bar{X}_{obs} = 12.5$

c. n = 1600:
   (i) $\bar{X}_{obs} = 12.1$
   (ii) $\bar{X}_{obs} = 12.2$
   (iii) $\bar{X}_{obs} = 12.6$

3. Tests (e.g., in Neyman-Pearson theory) are specified by giving a rule for when outcomes should be taken to reject H and accept J, and when not. A typical rule would be to reject H whenever the difference between the observed mean and the mean expected assuming H were true reaches a given small statistical significance level, e.g., .05, or .03, or .01. When the significance level that will lead to rejecting H is set out ahead of time at some fixed value, say .03, this value is called the size of the test.

4. Suppose the hypotheses being tested are as above:

   H: $\mu = 12$ vs. J: $\mu > 12.$

**Define (one-sided) test T+ (with size = .03):** Reject H and accept J whenever $\bar{X}_{obs}$ is statistically significantly different from 12 (in the positive direction) at level .03.

Equivalently we can define:
Test T+ (with size = .03): Reject H and accept J whenever $\bar{X}_{obs} \geq 12 + 2 \sigma_x$.

or
Test T+: Reject H and accept J whenever $D_{obs} \geq 2 \sigma_x$.

**Define (one-sided) test T+ (with size = .001):** Reject H and accept J whenever $\bar{X}_{obs}$ is statistically significantly different from 12 (in the positive direction) at level .001.

Equivalently we can define:
Test T+ (with size = .001): Reject H and accept J whenever $\bar{X}_{obs} \geq 12 + 3 \sigma_x$.

or
Test T+: Reject H and accept J whenever $D_{obs} \geq 3 \sigma_x$. 

2
5. Let \( \bar{X}^* \) be the **cut-off for rejection**.

Then \( \bar{X}^* \) for \( T^+ \) with size \( .03 = 12 + 2 \sigma \bar{X} \).

**Questions:** What is \( \bar{X}^* \) for this test when \( n = 100 \)?
(answer: 12.4).

What is \( \bar{X}^* \) for Test \( T^+ \) with size \( .001 \) and \( n = 100 \)?
(answer: 12.6.)

---

**Problem Set #2:**

1. Find \( \bar{X}^* \) for Test \( T^+ \) with size \( .03 \) and \( n = 25 \), and for \( n = 1600 \).
2. Which of the outcomes in Problem Set #1, a. and b., would lead to rejecting \( H \) at this level? The outcomes were:

   a. \( n = 25 \):
      (i) \( \bar{X}_{obs} = 12.6 \)
      (ii) \( \bar{X}_{obs} = 12.8 \)
      (iii) \( \bar{X}_{obs} = 13. \)

   c. \( n = 1600 \):
      (i) \( \bar{X}_{obs} = 12.1 \)
      (ii) \( \bar{X}_{obs} = 12.2 \)
      (iii) \( \bar{X}_{obs} = 12.6 \)

---

5. **The Type I error:** This is the error of rejecting \( H \) when actually \( J \) is true. The probability that test \( T^+ \) with size \( .03 \) commits a Type I error, written as

\[
\alpha = P(\ text {Test \ T^+ \ rejects \ H \ ; \ H \ is \ true}) = \\
= P( \bar{X} \geq \bar{X}^* \ ; \ H \ is \ true) = \\
= P( \bar{X} \ is \ statistically \ significant \ at \ level \ .03 | \ H \ is \ true) = .03
\]

So, the fixed size of the test ensures that the probability of a Type I error is no more than \(.03\).

**In general, by fixing a small size \( \alpha \), the test is assured of having a small probability, namely \( \alpha \), of committing a Type I error.**

6. Suppose Test \( T^+ \) with size \( \alpha = .03 \) rejects \( H \) and “passes” \( J \) with an outcome that just reaches the cut-off for rejection, i.e., suppose \( \bar{X}_{obs} = \bar{X}^* \).

We can show that the severity with which \( J \) passes test \( T^+ \) with \( \alpha = .03 = 1 - \alpha = 1 - .03 = .97 \).

By definition of severity, the severity with which a hypothesis \( J \) passes a test \( T \) with outcome \( e = P( \text{Test \ T \ would \ not \ yield \ such \ a \ passing \ result \ ; J \ is \ false} \)

\[
= 1 - P( \text{Test \ T^+ \ would \ yield \ such \ a \ passing \ result \ \mid \ J \ is \ false})
\]

Notice that “passing \( J \)” (that \( \mu > 12 \)) is the same as rejecting \( H \) in the case of test \( T^+ \). And “J is false” = \( H \) is true.

Therefore, this equals \( = 1 - P(\text{Test \ T^+ \ reject \ H \ ; \ H \ is \ true}) \)
and from #5 above, we have this is $= 1 - \alpha = 1 - .03 = .97$.

(Note: There is no difference in this test if we let $H: \mu \leq 12$)

7. **The Type II Error, $\beta$, and Power**

Once again, consider Test T+ (with size $\alpha = .03$) and $\bar{X}^*$ the cut-off for rejection of $H$ where

$$H: \mu = 12 \text{ vs. } J: \mu > 12.$$ 

**The Type II error** is the error of failing to reject $H$, even though $H$ is false and $J$ is true.  

The probability a test $T$ commits a Type II error, written as

$$\beta = P(\text{Test } T \text{ commits a Type II error}) = P(\text{Test } T \text{ accepts } H ; J \text{ is true})$$

where “accept $H$” just means you do not reject it. (We will clarify the precise way to interpret “accept $H$” as we proceed.)

Now, $J$ in test $T+$ is a **composite** or **complex** hypothesis. In order to calculate $\beta$, we have to consider particular point values of $\mu$ in the alternative $J$ (i.e., values of $\mu \geq 12$). Let $J'$ abbreviate a specific value of $\mu$ in the alternative $J$. Then

$$\beta = P(\text{Test } T+ \text{ accepts } H ; J') = 1 - P(\text{Test } T+ \text{ rejects } H ; J')$$

$$= P(\bar{X} < \bar{X}^* ; J') = 1 - P(\bar{X} \geq \bar{X}^* ; J')$$

**Case 1: $J' < \bar{X}^*$**

Draw a picture of the Normal curve, label $J'$ right at the midpoint, $H$ to the left of $J'$ and $\bar{X}^*$ somewhere to the right of $J'$. shade in the entire area under the curve to the left of $\bar{X}^*$. This shaded area, which must clearly be greater than .5, equals $\beta$. So, the probability of a Type II error in this case is not low, it is greater than .5.

**Case 2: $J' > \bar{X}^*$**
Draw a picture of the Normal curve, label \( J' \) right at the midpoint, \( \bar{X}^* \) to the left of \( J' \) and \( H \) to the left of \( \bar{X}^* \). Shade in the area under the curve to the left of \( \bar{X}^* \). This shaded area, which must clearly be less than .5, equals \( \beta \).

\[ H(12) \quad \bar{X}^* \quad J' \]

To Calculate a particular value for \( \beta \), i.e., \( P(\bar{X}^* < \bar{X} ; J'') \)

Note that it is always calculated using the cut-off value for rejecting \( H \), namely, \( \bar{X}^* \). Again, you begin by calculating a difference, but now, since the assumption is that \( J' \) is true, you calculate:

1. Find \( D = \bar{X}^* - J' \)
2. Divide \( D \) by \( \sigma_x \)
   (Note that in case 1, \( D \) is positive, in case 2, \( D \) is negative.)
3. Find the area to the left of the value you get in #2, under the standard Normal curve.
4. The area to the right of the value you get in #2 is the power of the test (to reject \( H \) and accept \( J' \)). (So power = 1 - \( \beta \))

Example: For \( T+ \) (with \( \alpha = .03 \)) and \( n = 100 \), the cut-off for rejecting \( H \), \( \bar{X}^* \), is 12.4. \( (\sigma_x = .2) \) Say you want to find \( \beta \) for \( J' \): \( \mu = 12.2 \).

You ask: What is the probability that Test \( T+ \) would accept \( H \) when in fact \( J' \) is true? That is: what is the probability that Test \( T+ \) would accept \( H: \mu = 12 \) when in fact the true value of \( \mu \) is 12.2?

To answer, you calculate \( D = 12.4 - 12.2 = .2 \) which is equal to 1 standard deviation. (Note, this is an example of case 1.)

The area to the left of 1 on the standard Normal curve is about .84. So \( \beta = .84 \).
Problem Set #3: For Test T+ (with $\alpha = .03$) and $n = 100$ with

$H: \mu = 12$ vs. $J: \mu > 12$ and $\sigma = 2$, so ($\sigma_x = .2$)

1. Find the probability that Test T+ commits a Type II error when $J': \mu = 12.6$.

A Useful Abbreviation: Since the probability of a Type II error, $\beta$, always varies with the particular point value of $J'$, it will be useful to abbreviate the probability that Test T+ commits a Type II error when $J': \mu = \mu'$ as simply $\beta(\mu')$ or $\beta(J')$.

So problem 1 in set #3 is to calculate $\beta(12.6)$.

I'll do this first one: $D = 12.4 - 12.6 = -.2$, which is $-1 \sigma_x$. The area to the left of -1 on the standard Normal curve = .5 - (the area between 0 and 1) = .5 - .34 = .16.

Note that this is an example of case 2: $J > \bar{X}$.*

2. Find $\beta(12)$.
3. Find $\beta(12.2)$
4. Find $\beta(12.4)$.

5. Find $\beta(12.7)$.
6. Find $\beta(12.8)$.
7. Find $\beta(12.9)$.
8. Find $\beta(13)$.

9. Plot these points on a curve on whose horizontal axis are values of $J'$ and whose vertical axis are the corresponding values of $\beta(J')$.

10. Explain why $\beta(H) = 1 - \alpha$.

11. Graph the corresponding values (in 1-8) for the power. This yields a power curve.

12. Explain: If Test T+ accepts $H$ (say it just misses the cut-off $\bar{X}*$) and $\beta(\mu')$ is very low, then “$\mu < \mu'$” passes a severe test.