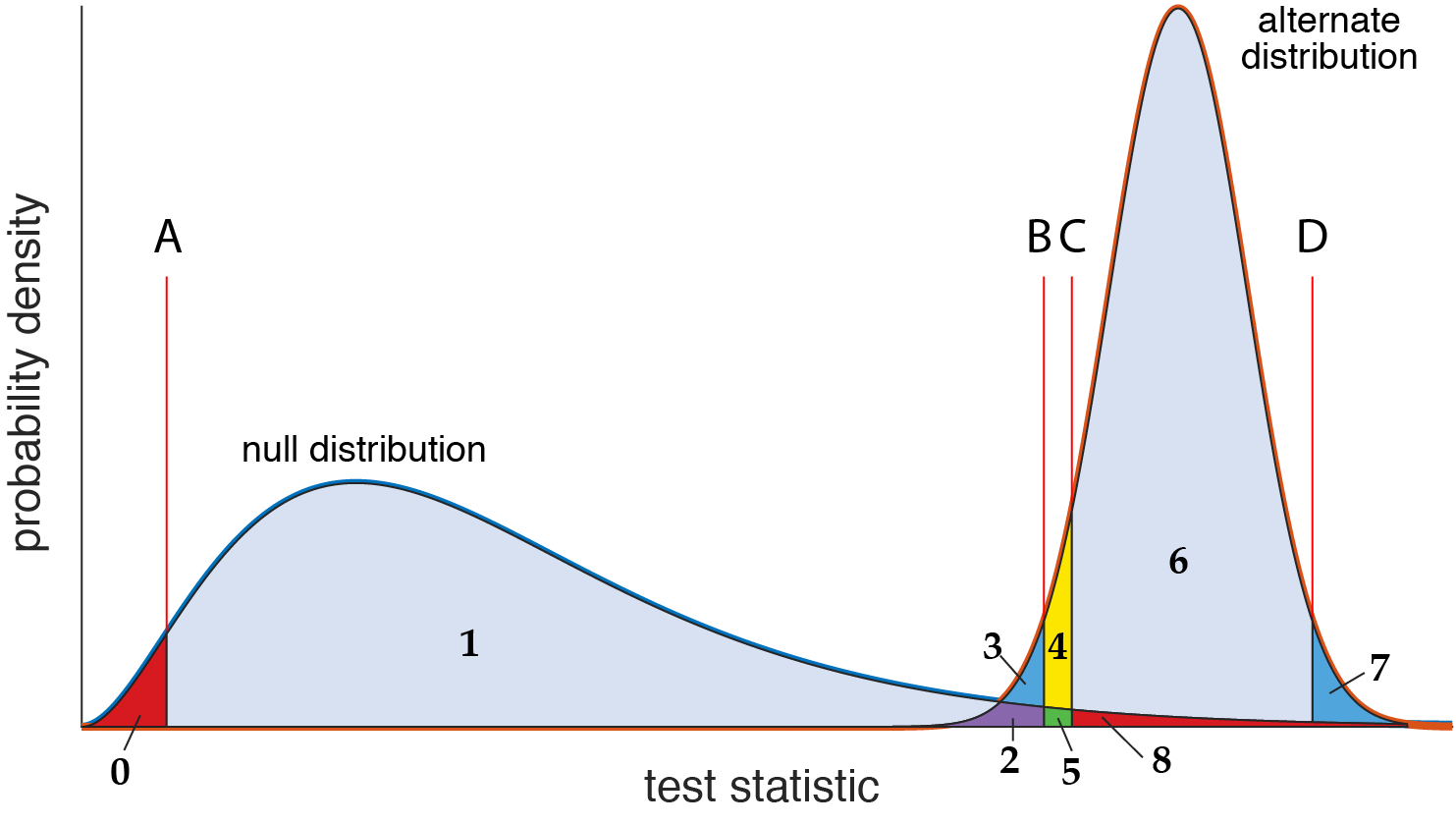
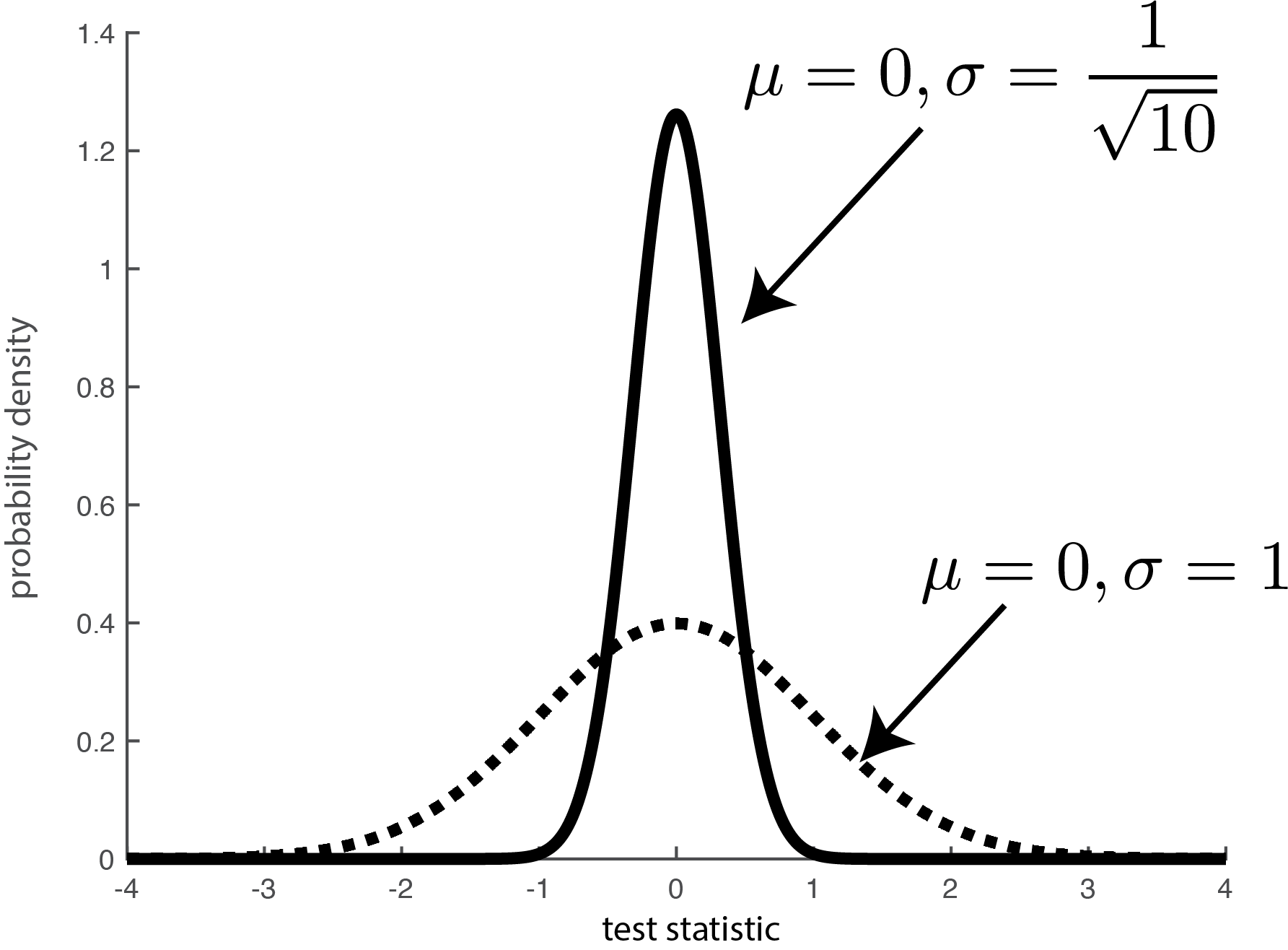
1. Indicate if each of the following statements about p-value is True or False:
   1. p-value is the probability that the observed value of the test statistic or more extreme happens under the null hypothesis.  
      True: this is the definition
   2. For a significance level of alpha, p-value always equals alpha.  
      False: p-value depends on data
   3. If p-value is larger than the significance level, the null hypothesis is accepted.  
      False: null hypothesis is never accepted
   4. If p-value is smaller than the significance level, the null hypothesis is rejected.  
      True: this is the definition of rejecting the null hypothesis
   5. p-value is the probability that the null hypothesis is false.  
      False: just wrong! NHT doesn’t let us talk about the probability of null hypothesis
   6. p-value is the probability that the alternate hypothesis is false.  
      False: just wrong! NHT doesn’t let us talk about the probability of alternate hypothesis either
   7. If p-value is smaller than the significance level, the result is said to be *statistically significant*.  
      True: that’s the definition of being ‘statistically significant’
2. In the figure below, vertical lines A and C indicates to the 2.5% tails of the null distribution, while B and D indicates the 2.5% tails of the alternate distribution. In other words, the red areas (area 0 and 8) sum to 5%. If the significance level alpha = 5%, which area corresponds to the probability of correctly rejecting null hypothesis when the alternate hypothesis is true (i.e., the statistical power)? (answer could be sum of multiple areas; assume a two-sided test)  
     
   
   1. Area 0 + Area 8
   2. Area 2 + Area 3
   3. Area 7
   4. Area 6 + Area 7 + Area 8
   5. Area 2 + Area 3 + Area 4 + Area 5
   6. Area 1 + Area 2 + Area 3
   7. Area 2 + Area 3 + Area 7

Incorrect! A null hypothesis test based on this null distribution will reject any test statistic less than A or greater than C. Now, recall that the **area under these probability density curves represents probabilities**. For example, Pr[ test statistic from the null distribution < A ] is simply the area under the curve of the null distribution, that is, the area 0. Now, if the alternate hypothesis is true, the probability of the test statistic corresponds to the area under the alternate distribution. For example, Pr[ test statistic from the alternate distribution < B ] is equal to the area 3 + area 2.  
  
The question is asking what is the probability that the test statistic from the alternate distribution will be rejected. We have already decided this happens when it is less than A or greater than C. It’s not going to be less than A, so really it is asking what is  
Pr[ test statistic from the alternate distribution > C ], which is area under the curve of the alternate distribution on the right side of C.

1. According to the central limit theorem, which of the following distributions tend to normal distribution? (choose all that applies)
   1. Binomial distribution as number of events (number of total coin flips) increase
   2. Mean of n independent samples from a chi-squared distribution as n increases
   3. Sum of m independent samples from a normal distribution as m increases
   4. Sampling distribution of estimated proportion assuming random sampling as the sample size increases
   5. Sampling distribution of the mean from ANY population distribution as the sample size increases
   6. All of the above

Incorrect! The central limit theorem applies to sampling distribution of the mean, or sum of many random samples from a fixed distribution.

1. Consider two independent random samples of size *N* and *M* from the same population. Figure below shows the sampling distribution of the mean corresponding to the two different sample sizes. These sampling distributions are both normal distributions with parameters shown in the figure. Suppose *N* is much greater than *M*.



* 1. How much larger is the *N* relative to *M*? (Hint: what is the formula for the standard deviation of the sampling distribution of the mean?)

10 [ exact numerical answer ]  
  
Incorrect! Recall that the standard deviation of the sample mean is given by  
 ../../../../../../../var/folders/t3/xxwwsrz93yl3ml3h5rl9k3zc0000gn/T/LaTeXiT-2.

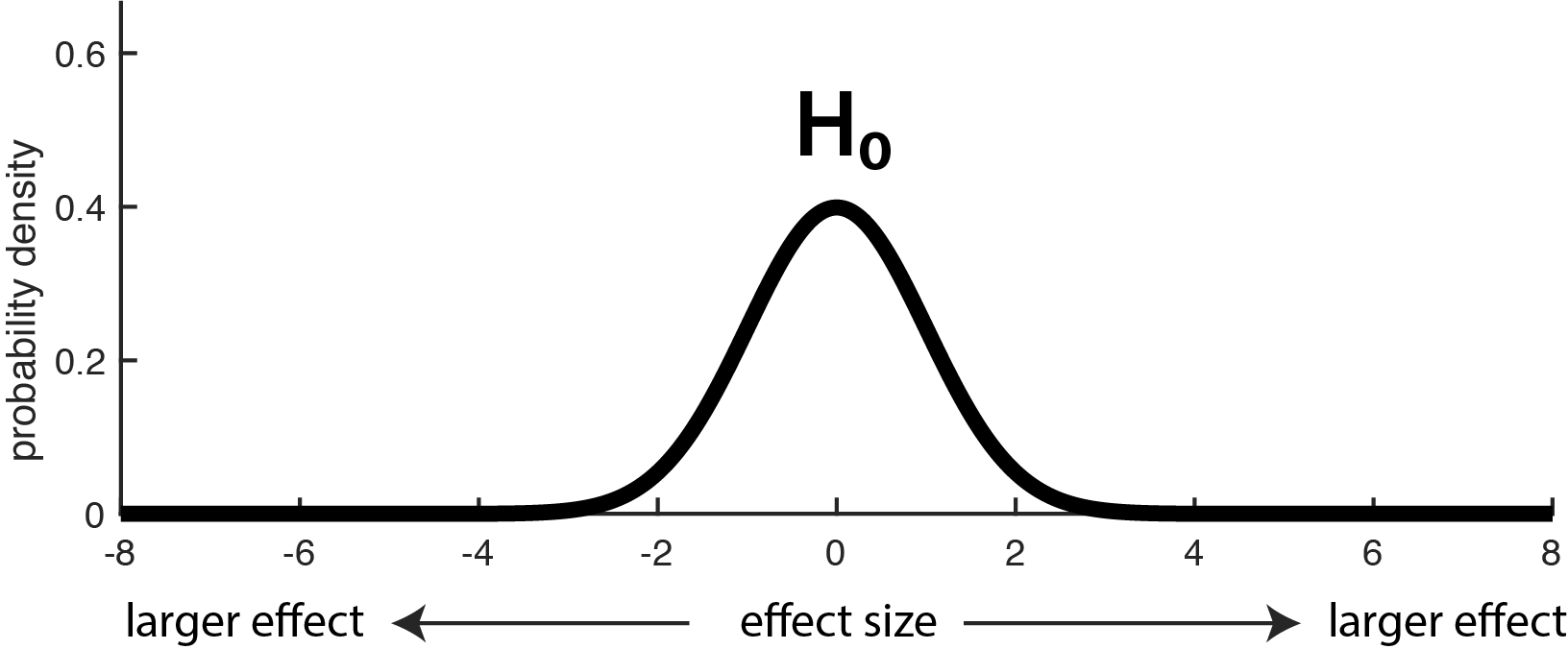
where ../../../../../../../var/folders/t3/xxwwsrz93yl3ml3h5rl9k3zc0000gn/T/LaTeXiT-2. denotes the standard deviation of the population, and *n* denotes the sample size. Importantly, the ../../../../../../../var/folders/t3/xxwwsrz93yl3ml3h5rl9k3zc0000gn/T/LaTeXiT-2. in the plot is the standard devation of the sample mean, which in this equation is ../../../../../../../var/folders/t3/xxwwsrz93yl3ml3h5rl9k3zc0000gn/T/LaTeXiT-2..

This equation tells you that as your sample size gets larger for estimating the mean, the standard devaition of the mean gets smaller. In other words, your sampling distribution concentrates around the mean, and gets tighter.

More specifically, the plot shows that the standard deviation became smaller by a factor of ../../../../../../../var/folders/t3/xxwwsrz93yl3ml3h5rl9k3zc0000gn/T/LaTeXiT-2.. Therefore, we can infer that N = 10 \* M, that is, **the sample size N was 10 times larger than samples size M**.

* 1. In the context of null hypothesis testing, these two distributions can correspond to two different null distributions where one experiment has larger number of participants compared to the other. Suppose the observed test statistic from the experiment yielded -2. How does the p-value associated with this observed test statistic change as the sample size increased from *M* to *N*?
     1. Increases
     2. Decreases
     3. Remains constant
     4. Not predictable  
          
        Incorrect! p-value is the area under the tail times 2 (i.e., probability of getting the observed test statistic -2 or more extreme). The tail for N is clearly thinner thanks to the smaller standard deviation. Therefore, the p-value decreases when sample size increases to N.

1. Effect size is the deviation from the expected test statistic under the null hypothesis. (see figure below). In other words, this null hypothesis test is designed to (statistically) decide if the effect size of the experimental outcome is non-zero. Usually the effect size has a scientific (biological) meaning–for example, the effect could be how different your control group is from the treatment group. Does **smaller p-value** imply about the effect size? (Note that the null distribution can change for the same test statistic as a function of sample size. Hint: see problem #4)
   1. Effect size tends to increase
   2. Effect size tends to decreases
   3. Effect size tends to remain constant
   4. Effect size is not necessarily related to p-value



See explanation of 4b. For the same effect size, say 2, as the number of samples increases the p-value decreases. Therefore, smaller p-value is not necessarily an indication of a large effect size, but it could simply be the result of an increased sample size.