Customer C Spend on Order $\mathbf{X} = c * r * t + \epsilon$

$c =$ Customer's mean order amount in dollars where $c \sim \mathcal{N}(\$100, \$10)$.

$r =$ Scalar corresponding to mean retailer order amount where $r \sim \mathcal{N}(1.0, 0.05)$.

$t =$ Whether or not the customer recieved the experimental treatment $t \in [1, 1.1]$.

$\epsilon =$ Random noise where $\epsilon \sim \mathcal{N}(\$10, \$1)$.

Recall that the data is clustered (we aspire for customers to place multiple orders!), thus:

The number of orders placed by Customer C $\sim exp(\lambda = 1)$.

If we were implementing a difference of means t-test, then we could simply apply the formula below to get the appropriate sample size.

$$n = 2\left(\frac{Z_{1-(\alpha/2)} + Z_{1-\beta}}{ES}\right)^2 \text{ where } ES = \frac{|\mu_1 = \mu_2|}{\sigma}$$

Here, α is the probability of a Type I error (false rejection of the NULL hypothesus, i.e. a false positive) and β is the probability of a Type II error (i.e. a false negative). Observe how if we want to make our test more rigorous and reduce the false positive rate, the the left side of the numerator increase and the resulting recommended sample size - n, also increases. Similarly, if we want to reduce the probability for failing to detect the difference across groups, then we will need to increase the right-hand side of the numerator - also increasing the recommended sample size.
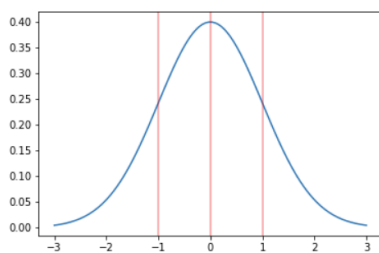
$$ES = \frac{|\mu_1 = \mu_2|}{\sigma}$$
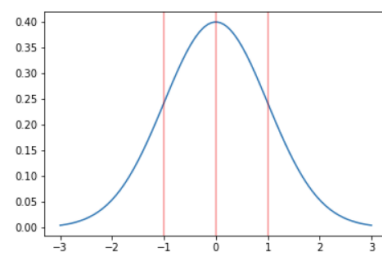
$$y = \beta_0 + X_1\beta_1$$
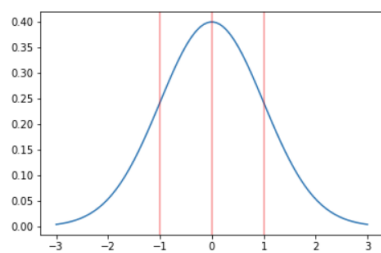
$$y = \beta_0 X_1 + ... + \beta_n X_n + \epsilon$$
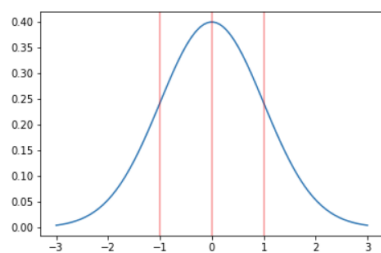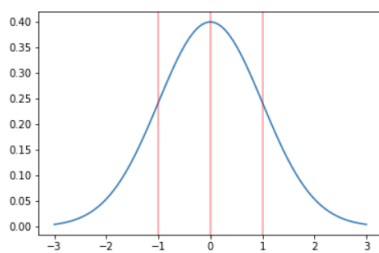
(a) μ +/- $[1]\sigma$
with 500 iterations.
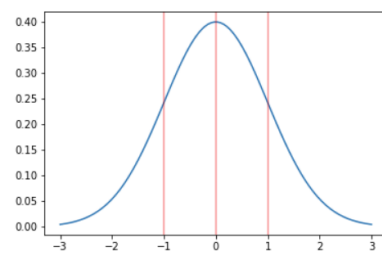
(d) μ +/- $[1]\sigma$
with 1,000 iterations.

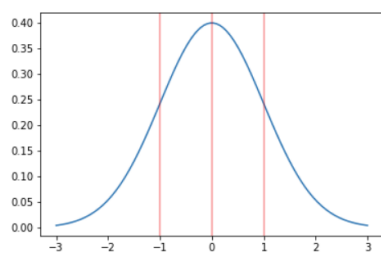(g) μ +/- $[1]\sigma$
with 2,000 iterations.

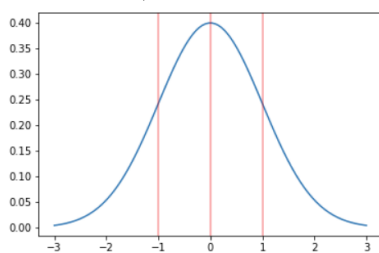(b) μ +/- $[1,2]\sigma$
with 500 iterations.

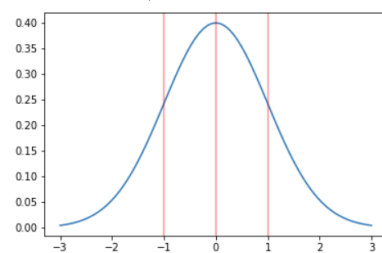(e) μ +/- $[1,2]\sigma$
with 1,000 iterations.

(h) μ +/- $[1,2]\sigma$
with 2,000 iterations.

(c) μ +/- $[1,2,3]\sigma$
with 500 iterations.

(f) μ +/- $[1,2,3]\sigma$
with 1,000 iterations.

(i) μ +/- $[1,2,3]\sigma$
with 2,000 iterations.

Figure 1