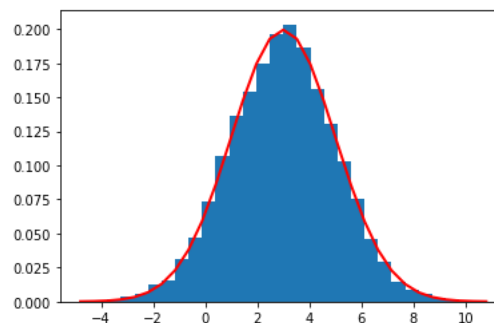


## Testing Hypothesis in Python

- Generate a normally distributed data, with mean=3 and standard deviation=2, for 10,000 observations:

```
import numpy as np
mu, sigma = 3, 2
s = np.random.normal(mu, sigma, 10000)
```

- Plot the normal distribution overlaid on top of a histogram in matplotlib so you get the following result:



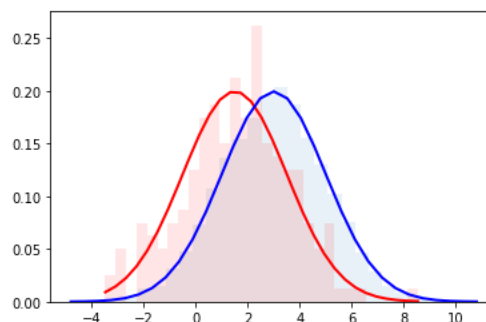
- Create a sub-sample of this population with mean of 1.5, a sigma of 2.0, and 200 observations.

```
import numpy as np
sample_mean, sample_sigma = 1.5, 2
sample = np.random.normal(sample_mean, sample_sigma, 200)
```

-Use:

```
count, bins, ignored = plt.hist(s, 30, alpha=0.1, density=True)
```

- Compare both distributions on a chart. The chart should look like this:

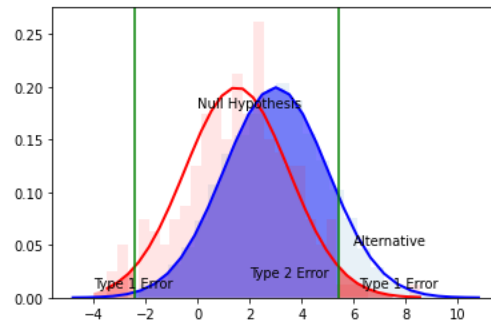


- Use red for the sample distribution and blue for the real population distribution.
- Use 'from scipy.stats import norm' to create 95% confidence intervals. Consider the interval as

```
plt.axvline(ci[0], color='g')
```

```
plt.axvline(ci[1],color='g') z
```

-Create a graph that shows the null and hypotheses, the type 1 and 2 errors that looks like this:



-Compute the z-score, which is equal to  $(\text{sample mean} - \mu) / \sigma$ . What is the p-value and do you reject it?