hw3

1.

(a) Null hypothesis: There's no difference between Caffeine group and Sleep group in average recall ability. Alternative hypothsis: There is difference between Caffeine group and Sleep group in average recall ability.

Since we assume equal population variance, $SE_{y1-y2} = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(15-1)3.3^2+(12-1)3.6^2}{15+12-2}} = 3.435$,

Test statistics $T = \frac{\overline{y1}-\overline{y2}}{SE_{y1-y2}} = \frac{15.25-12.25}{3.435} = 0.873$

$SE_1^2 = s_1^2/n_1 = 3.3^2/15 = 0.726$

$SE_2^2 = s_2^2/n_2 = 3.6^2/12 = 1.08$

Degree of freedom $df = \frac{(SE_1^2+SE_2^2)^2}{SE_1^4/(n_1-1)+SE_2^4/(n_2-1)} \approx 23$

At 95% confidence level, $t_{23}^{0.975} = 2.069$.

Since $T = 0.873 < 2.069$, we fail to reject the null hypothesis. We are not justified in concluding that caffeine group and sleep group have different average recall ability.

(b) If we assume different population variance, $SE_{y1-y2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 1.344$

Test statistics $T = \frac{\overline{y1}-\overline{y2}}{SE_{y1-y2}} = 2.23$

Degree of freedom $df = n_1 + n_2 - 2 = 25$

At 95% confidence level, $t_{25}^{0.975} = 2.060$

Since $T = 2.23 > 2.060$, we reject the null hypothesis, and accept the alternative hypothesis.

(c)

i. $T = \frac{\overline{y}-\mu}{s/\sqrt{n}} \sim Student_{n-1}$

$P(-t_{n-1}^{1-\alpha/2} \times s/\sqrt{n} \leq \overline{y} \leq t_{n-1}^{1-\alpha/2} \times s/\sqrt{n}) = 1-\alpha$, $(\alpha = 0.05)$

For sleep group $t_{n-1}^{1-\alpha/2} = t_{14}^{0.975} = 2.145$

$CI_{sleep} = [15.25 \pm 2.145 \times 3.3/\sqrt{15}] = [15.25 \pm 1.83] = [13.42, 17.08]$

For caffeine group $t_{n-1}^{1-\alpha/2} = t_{11}^{0.975} = 2.201$

$CI_{caffeine} = [15.25 \pm 2.201 \times 3.6/\sqrt{12}] = [12.25 \pm 2.29] = [9.96, 14.54]$

Interpretation: The probability that the average recall ability of sleep group lies in the range [13.42, 17.08] is 95%. The probability that the average recall ability of caffeine group lies in the range [9.96, 14.54] is 95%. The confience intervals of the two groups are not completely separate, they overlap a little.

ii. Confidence interval is an estimate of the range of the average recall ability of each group, it's not a point estimate or a percentage estimate, 95 % confidence inerval means the probability that the average recall ability of the group within the interval range is 95%.

iii. $T = \frac{\overline{y_1}-\overline{y_2}-\delta}{SE_{y1-y2}} \sim Student_{n_1+n_2-2}$

$SE_{y1-y2} = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = 1.344$

$t_{n_1+n_2-2}^{1-\alpha/2} = t_{25}^{0.975} = 2.060$

$P(-t_{n_1+n_2-2}^{1-\alpha/2} \times SE_{y1-y2} \leq \overline{y_1} - \overline{y_2} - \delta \leq t_{n_1+n_2-2}^{1-\alpha/2} \times SE_{y1-y2}) = 1-\alpha$, $(\alpha = 0.05)$

$-2.060 \times 1.344 \leq 15.25 - 12.25 - \delta \leq 2.060 \times 1.344$

$\delta \in [0.23, 5.77]$

(d) When using the hypothesis testing approach, we either reject a hypothesis or fail to reject a hypothesis based on alpha, which is a predefined hard limit. But confidence interval provides a range estimate of the mean, and it contains more information.

2.

(a) Suppose random variable Y denotes whether or not a random responder supports same-sex marriage. Y folows Bernoullie distribution with $\mu = p$ and $\sigma^2 = p(1-p)$ According to Central Limit Theorem, $\frac{\overline{y_n}-p}{\sqrt{p(1-p)/n}} \sim N(0,1)$, p=0.6, n=1024, 90% confidence interval is $p \pm z^{1-10\%/2}\sqrt{p(1-p)/n} = [0.575, 0.625]$, 95% confidence interval is $p \pm z^{1-5\%/2}\sqrt{p(1-p)/n} = [0.57, 0.63]$ 95% confidence interal is wider, since the confidence is higher, which means the probability of the proportion in the population who favor legalizing same-sex marriage within the interval range is higher, the interval needs to be wider.

(b) We want to calclate the minimum sample size that makes 0.55 outside the 95% confidence interval. So $p - z^{1-5\%/2}\sqrt{p(1-p)/n} \geq 0.55$, p=0.6, $z^{1-5\%/2} = 1.96$, $n \geq 369$, the minimum number of 2015 interviews is 369.

3.

(a) Null hypothesis is the same as alternative hypothesis.

(b) It's wrong to use independent two-sample t-test, since the two samples are not independent, 100 students contains the 56 males.

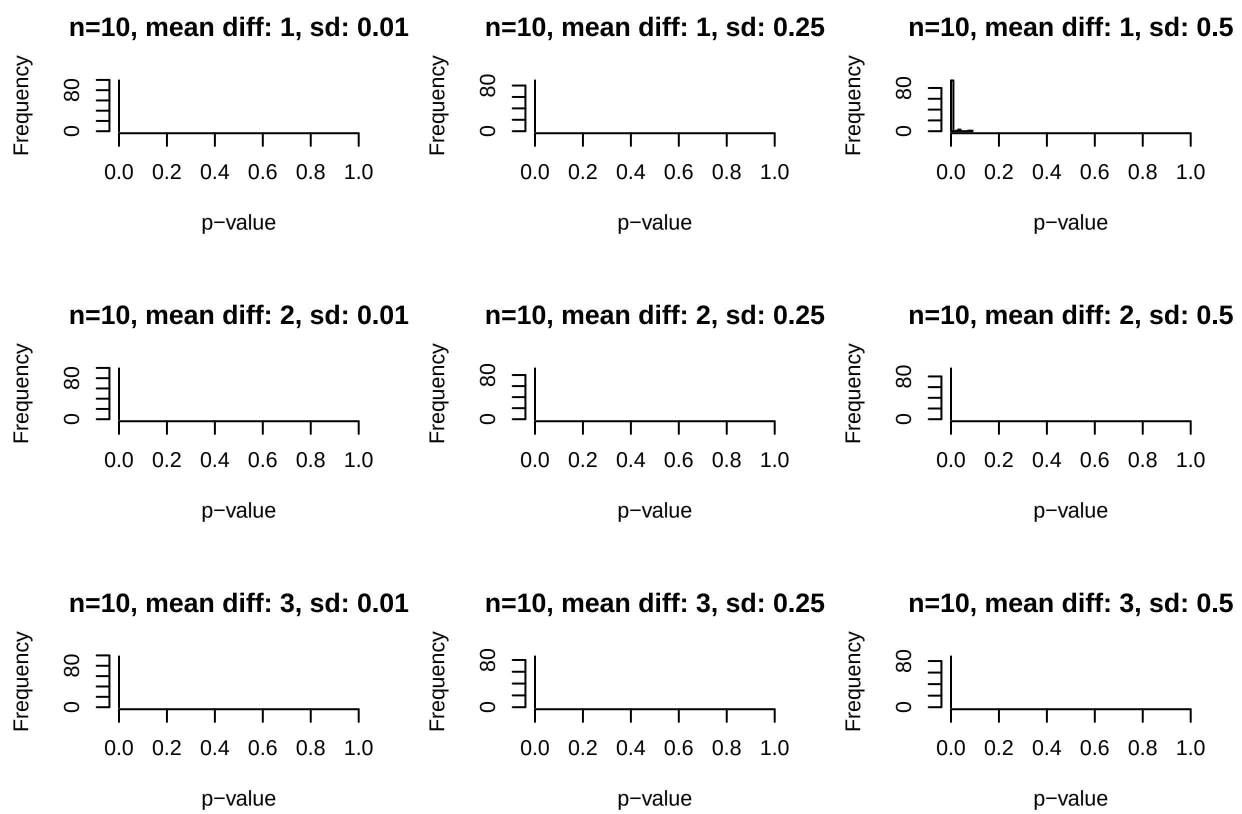
(c) First, we don't know alpha (the predefined significance level), so we don't know whether we should reject the null hypothesis. In addition, p-value=0.94 is too high, assuming a commonly chosen significance level (0.05), we should fail to reject the null hypothesis.

(d) Since we don't know whether the distribution is symmetrical or not, we can't assume the p-value for the one-sided test is half of the p-value for the two-sided test.

4.

(a) & (b)

```r
ttest.pval <- function(mean, sd, n){
  sample1 <- rnorm(n, mean, sd)
  sample2 <- rnorm(n, 0, sd)
  result <- t.test(sample1, sample2)
  return(result$p.value)
}
means <- c(1, 2, 3)
sds <- c(0.01, 0.25, 0.5)
sizes <- c(10)
par(mfrow=c(3,3))
for(size in sizes){
  for(mean in means){
    for(sd in sds) {
      pvals <- replicate(100, ttest.pval(mean, sd, size))
      hist(pvals, main=paste0("n=", size, ", mean diff: ", mean, ", sd: ", sd), xlab="p-value", xlim=c(
    }
  }
}
```



(c)

The larger the difference of the means, the easier to detect the differences. The larger the standard deviation, the harder to detect the differences. The larger the sample, the easier to detect the differences.