Stat 135, Fall 2006 A. Adhikari

HOMEWORK 6 SOLUTIONS

1a. Under the null hypothesis X has the binomial (100, .5) distribution with $E(X) = 50$ and $SE(X) = 5$. So $P(|X - 50| > 10)$ is (approximately) two tails of the standard normal beyond the points $-z = (39.5 - 50)/5 = -2.1$ and $z = (60.5 - 50)/5) = 2.1$. That's 3.58%.

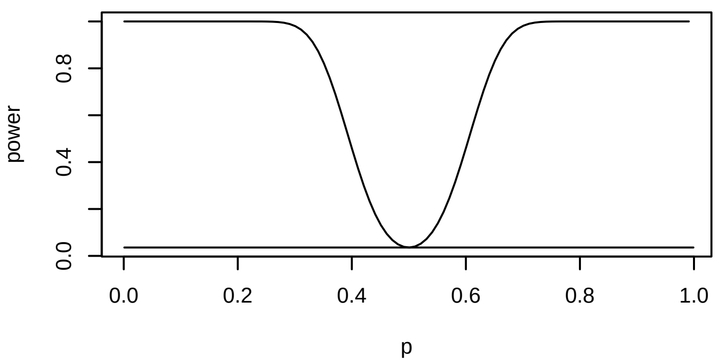
I know the "answer" in the book is 4.56%, but that's a normal approximation without the continuity correction. You know better than that. R computes the exact binomial probability of rejection to be 3.52%, so 3.58% is a better approximation than 4.56%.

b. The rejection region is: X greater than 60 or less than 40. The distribution of X is approximately normal with mean $100p$ and SE $\sqrt{100p(1-p)}$. So for each fixed p the power is going to be the sum of two normal tails, beyond the points 40.5 and 60.5. I used the following code.

```
> p ← seq(.001, .999, by = .01)
> power ← pnorm(39.5, 100*p, sqrt(100*p*(1-p))) + 1 - pnorm(60.5, 100*p, sqrt(100*p*(1-p))
> plot(p, power, type="l")
> lines(c(.001, .999), c(.0358, .0358))
```

The last command plots the horizontal line at the level $\alpha = 0.0358$. It's clear that the power converges to the level as p converges to 0.5, and that it converges to 1 as p converges to 0 or 1.

The horizontal line is at the level alpha = 3.58%.

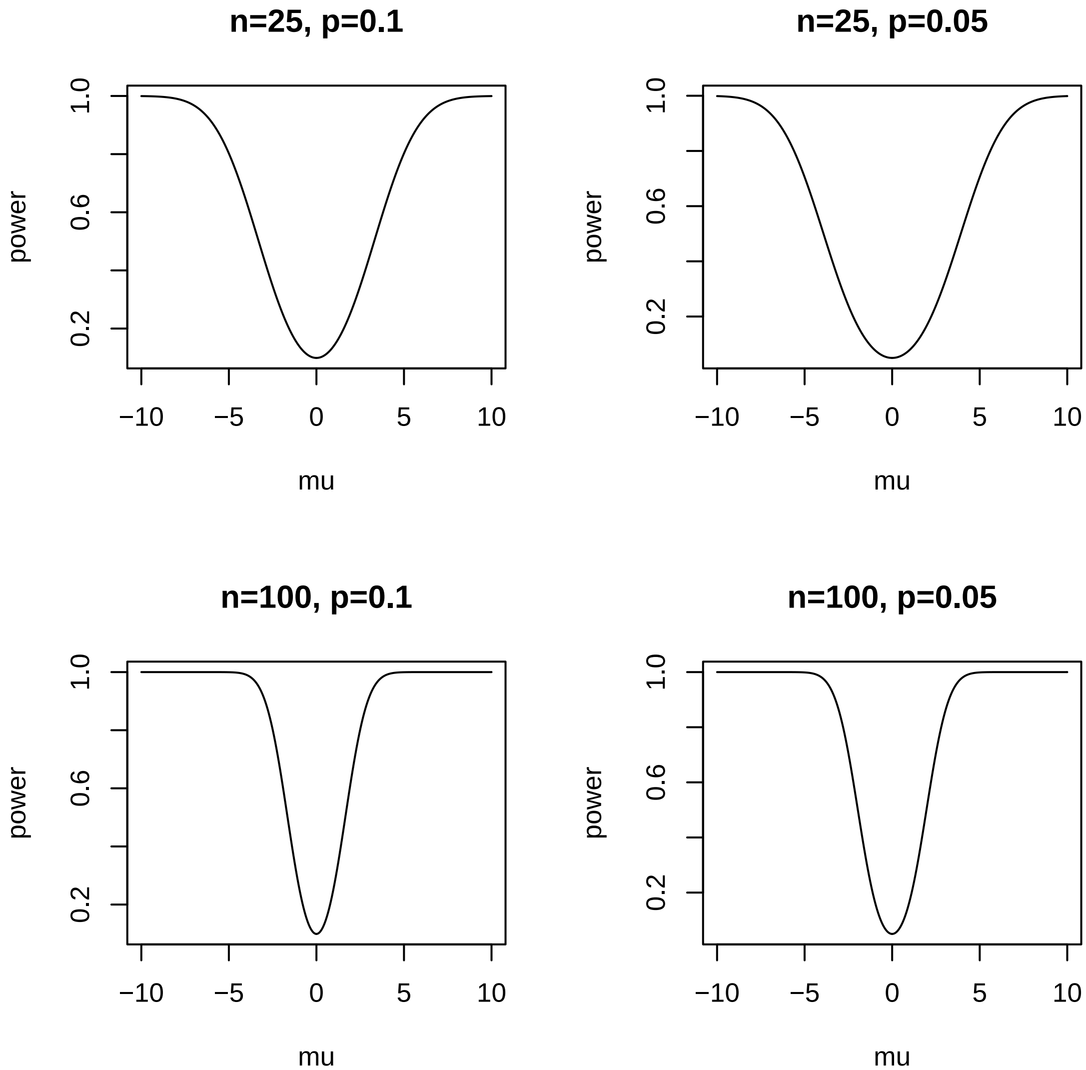


2. We derived the relevant two-tailed test in lecture. It is a two-tailed test which rejects when $\bar{X}$ is far from 0 on either side. When $n = 25$, the distribution of $\bar{X}$ is normal with mean μ and SE $10/\sqrt{25} = 2$. Therefore under H_0, the distribution of $\bar{X}$ is normal with mean 0 and SE 2. If the significance level is $\alpha = 0.1$ the critical z is 1.65, so the rejection region is "$\bar{X} < -3.3$ or $\bar{X} > 3.3$. For a fixed value μ in the alternative, the power is the total area to the left of -3.3 and to the right of 3.3, under the normal curve with mean μ and SE 2. I chose to plot power for μ in the range $(-10, 10)$ because that's a few SEs on either side of 0.

```
> mu ← seq(-10, 10, by=0.1)
> power ← pnorm(-3.3, mu, 2) + 1 - pnorm(3.3, mu, 2)
```

If $\alpha = 0.05$ the critical z is 1.96 so the rejection region is "$\bar{X} < -3.92$ or $\bar{X} > 3.92$.

When $n = 100$ the disribution of $\bar{X}$ is normal with mean μ and SE 1. For $\alpha = 0.1$ the rejection region is "$\bar{X} < -1.65$ or $\bar{X} > 1.65$". For $\alpha = 0.05$ the rejection region is "$\bar{X} < -1.96$ or $\bar{X} > 1.96$". The power against any μ in the alternative is the chance of the rejection region computed using the normal curve with mean μ and SE 1.



You can see that for fixed n any fixed value of μ in the alternative, the power is less when α is small. That makes sense because if you allow yourself less Type I error you should expect to make more Type II error and hence have smaller power. It is also clear from the plots that for fixed α the test based on the larger sample is more powerful.

3. Since the MLE of θ is $1/\bar{X}$, the generalized likelihood ratio is

$$\frac{\prod_{i=1}^n \theta_0 e^{-\theta_0 X_i}}{\prod_{i=1}^n \frac{1}{\bar{X}} e^{-X_i/\bar{X}}} \quad = \quad [\theta_0 e \bar{X} e^{-\theta_0 \bar{X}}]^n$$

This is small when $\bar{X}e^{-\theta_0 \bar{X}}$ is small.

4a. Use the result of the previous problem. The function $g(x) = xe^{-\theta_0 x}$ is small when x is very large (because $e^{-\theta_0 x}$ gets small at a much faster rate than x gets large) and when x is very small (because $e^{-\theta_0 x}$ gets close to 1). That explains the form of the rejection region.

b. Under $\theta_0 = 1$, the given chance is the significance level.

c. $\sum X_i$ has the gamma (10, 1) distribution because it is the sum of 10 i.i.d. exponentials with rate 1. So $\bar{X}$ is 1/10 times a gamma random variable. Multiplying a gamma random variable by a constant yields another gamma, with a change only in the scale parameter. So $\bar{X}$ has the gamma (10, λ) density where λ can be found by equating means:

$$1 = E(\bar{X}) = \frac{10}{\lambda}$$

so $\lambda = 1$. Use this knowledge and R to find the cutoffs x_0 and x_1. Those should be respectively the 2.5th and 97.5th percentiles of the gamma (10, 10) distribution, which are about 0.48 and 1.71 according to R.

d. If you forget the theory, simulate. Draw 1000 i.i.d. samples, each consisting of 10 i.i.d. exponential (1) observations. Compute the mean of each sample, and use as your cutoffs the 2.5th and 97.5th quantiles of the observed distribution of means.

5. Both the hypotheses are simple, so by the Neyman-Pearson Lemma the likelihood ratio test is the most powerful. So the problem amounts to finding the power of the likelihood ratio test.

The likelihood ratio test rejects when $1/2X$ is small, that is, when X is large. The null distribution of X is uniform[0,1], so to achieve the level $\alpha = 0.1$, the test must reject when $X > 0.9$.

The power is the chance of the rejection region under the alternative, so it is equal to

$$\int_{0.9}^{1} 2x dx = 1^2 - 0.9^2 = 0.19$$

6a. Because the maximum likelihood estimate of p is X/n, the generalized likelihood ratio is

$$\frac{\binom{n}{X} 0.5^n}{\binom{n}{X}(X/n)^X(1-X/n)^{n-X}} = \frac{0.5^n}{(X/n)^X(1-X/n)^{n-X}}$$

b. The numerator in the ratio above does not depend on X, so the ratio will be small when the denominator is small. The denominator is

$$(X/n)^X(1-X/n)^{n-X} = \frac{1}{n^n}X^X(n-X)^{n-X}$$

This is large when $g(X) = X^X(n-X)^{n-X}$ is large. The function g is a function on $[0, n]$. It is symmetric about $n/2$ because $g(u) = g(n-u)$. And g is smallest at $n/2$, rising symmetrically on either side. If you're not convinced, compute a few values, plot g, or find its minimum.

So $X^X(n-X)^{n-X}$ is large when X is far from $n/2$, i.e., if $|X - n/2|$ is large.

c. The null distribution of X is binomial $(n, 0.5)$ which is a symmetric distribution. The rejection region is going to be two equal tails of the binomial, with the area of each tail equal to half the given significance level. For large n the binomial can be approximated by the normal with mean $n/2$ and SE $0.5\sqrt{n}$.

d. The rejection region is: $X < 3$ or $X > 7$. The significance level is $\alpha = 0.109375$, obtained by the R command

> 2*pbinom(2, 10, 0.5)

e. This is exactly the same problem as **1b** above, where you got the normal approximation of α. Use `pbinom` to check that the exact level is 0.0352, as discussed in the answer to **1b**.

7a. R says the sample means are 0.554625 and 1.62398, so those are my estimates. The estimate of the difference is 1.07, roughly. Yes, you can say it is 1.069355 or −1.069355 if you like.

b. The pooled estimate for the variance, 0.7666932. My lists were called `x` and `y` so I did

> (3*var(x) + 4*var(y))/7

c. $s_p\sqrt{1/4+1/5} = 0.587$, roughly.

d. Use the upper 5% point of the t_7 distribution, which is 1.895. The interval is $1.07 \pm 1.895 \times 0.587$, which is (-0.0424, 2.18).

e. I'll accept either answer. An argument for the one-sided test is that you're already told that the means may be different so it's the direction of the difference that's more interesting. An argument for the two-sided test is that you're told that the means may be the same, so you should try to find out whether they were the same or not.

f. The t-statistic is $(1.07-0)/.587 = 1.82$, roughly. The two-sided p-value is 0.11, roughly. I did

> 2*(1 - pt(1.82, 7))

g. No, but only just no. Because we're doing a 10%-level test, this is consistent with the fact that the 90%-confidence interval in **d** contains 0, but only just.

h. The variance of the difference would be exactly 0.670824 (use the known σ instead of s_p) and you would use the normal curve instead of the t. There's no need to do it all again.

8. Under H_0, $\bar{X}-\bar{Y}$ has the normal distribution with mean 0 and variance $10 \times \sqrt{2/n}$. The critical value is $z = 1.3$, so the rejection region is $\bar{X} > 13\sqrt{2/n}$. Under the particular alternative in the question, the distribution of $\bar{X}-\bar{Y}$ is normal with mean 2. The power is the area to the right of $13\sqrt{2/n}$ under this curve, and the area has to be 0.5. So $13\sqrt{2/n}$ has to be the mean of the curve.

Therefore $13\sqrt{2/n} = 2$, so $n = 85$.

9. The issue of statistical significance doesn't arise. There's nothing random, no probability model of any kind. The inner planets are simply more dense, on average, than the outer planets. There's no chance involved.

10a. Yes, it makes sense. The 55% is a percent from all Berkeley freshmen so there's no randomness there, but the randomness is in the fact that the 61% comes from a random sample. The test, therefore, is about the unknown proportion p of 2004 freshmen in the nation who would disagree with Statement B.

$H_0 : p = 0.55$, same as Berkeley. The sample percent is higher due to chance.

$H_A : p > 0.55$. The high sample percent reflects a high population percent. [It's also fine to say $H_A : p \neq 0.55$.]

Under the null, $\hat{p}$ has approximately normal distribution with mean 0.55 and standard error $\sqrt{.55 \times .45/1000} = 0.0157$. On this scale, .61 corresponds to $z = 3.82$ so the p-value is tiny and the null is rejected. The difference is statistically significant (indeed, highly statistically significant).

b. The question doesn't make sense because neither of the percents is from a random sample. Those two percents are just different. No p-values are needed, nor are they appropriate.

c. The question makes sense because both percents come from random samples. But because they come from the same random sample, they are dependent. In order to answer the question it is necessary to have some way of measuring the dependence. That can't be done based on the information given.

You should note, however, that the size of the difference is very great: 28%. It's hard to imagine a reasonable measure of dependence that would not make such a big difference statistically significant. Still, you can't do a test with the information given.

d. Both percents are from random samples, so the question makes sense. The test compares the national percent of freshmen agreeing with Statement A, in 1990 and 2004. The data come from two independent random samples.

$H_0 : p_{90} = p_{04}$ versus $H_A : p_{90} < p_{04}$ (yes, you can say $\neq$ instead of $<$)

Under H_0, $\hat{p}_{90} - \hat{p}_{04}$ has approximately normal distribution with mean 0. To compute the SE, first compute the pooled estimate of the common population proportion, $.5 \times .22 + .5 \times .33 = .275$. The SE of $\hat{p}_{90} - \hat{p}_{04}$ is approximately $\sqrt{0.275 \times (1 - 0.275) \times (1/1000 + 1/1000)} = 0.02$. So the difference in proportions is expected to be 0 with an SE of about 0.02. The observed difference was 0.11. That's more than 5 SEs away from the expected 0, so the difference is statistically significant. The data indicate that the two population proportions are not the same.