MTH 375
Fall 2022
Hw 10 key

1. $H_0: \ p_1 = \cdots = p_6 = 1/6$ vs. H_A : Not so.

number	1	2	3	4	5	6
frequency	20	15	17	11	18	19

We use the χ^2 test for goodness of fit. Under H_0, o_i is given by the table, and $e_i = 100/6$ for $i = 1, 2, 3, 4, 5, 6$. Using R,

```
> qchisq(df=5,.95)
[1] 11.0705
> c(20,15,17,11,18,19)->o
> c(100/6,100/6,100/6,100/6,100/6,100/6)->e
> sum((o-e)^2/e)
[1] 3.2
```

we see that if X is χ^2 with 5 df then $P(X > 11.07) = \alpha = .05$, so the appropriate test is "Reject H_0 if $X > 11.07$."

R also tells us that, for the given data, $x = \sum_i (e_i - o_i)^2/o_i = 3.2$. We do not reject H_0; the data is consistent with a fair die.

2. Let $f(x, \theta)$ be the pdf of the population the given data is sampled from.

$H_0: \ f(x; \theta) = \frac{1}{\theta} e^{-x/\theta}, \ x > 0$ vs. H_A: Not so.

We use the χ^2 test. The MLE $\hat{\theta} = \overline{X} = 176.0/60$. Our hypothesized pdf is $\exp(\theta = 176/60)$.

The groups for this test are determined by breaking the domain of f into 5 intervals of equal probability.

```
> qexp(rate=60/176,c(.2,.4,.6,.8))
[1] 0.6545544 1.4984218 2.6877861 4.7210179
```

(As I explained in class, in R notation, rate$= 1/\theta$.)

The intervals are (0,0.65), (0.65, 1.50), (1.50, 2.69), (2.69, 4.72), (4.72,∞). Under H_0, $e_i = 60 \cdot .2 = 12$, for $i = 1, 2, 3, 4, 5$. Counting shows that $o_1 = 11, o_2 = 15, o_3 = 8, o_4 = 11, o_5 = 15$.

```
 > qchisq(df=3,.95)
[1] 7.814728
> c(11,15,8,11,15)->o
> c(12,12,12,12,12)->e
> sum((o-e)^2/e)
[1] 3
```

Since we have estimated the parameter θ, our χ^2 statistic has $5-2=3$ df. R tells us that if X is χ^2 with 3 df then $P(X > 7.81) = \alpha = .05$, so the appropriate test is "Reject H_0 if 7.81."

R also tells us that, for the given data, $x = \sum_i (e_i - o_i)^2/o_i = 3$. We do not reject H_0; the data is consistent with an exponential pdf.

3. Find the equation of the line that best fits the points

$$\{(-4,5),(-4,4),(-3,2),(-2,3),(-2,1),(0,1),(0,2),(1,3),(1,0),(2,0),(2,-1),(3,1)\}$$

(in the sense of least squares). Sketch a graph containing these points and your line.

```
> c(-4,-4,-3,-2,-2,0,0,1,1,2,2,3)->x
> c(5,4,2,3,1,1,2,3,0,0,-1,1)->y
> lm(y~x)->line
> summary(line)

Call:
lm(formula = y ~ x)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.4769     0.3599   4.103  0.00213 **
x            -0.5462     0.1512  -3.612  0.00475 **
```

R tells us that the required line is $y = 1.48 - 0.55x$.

4. Test the hypotheses $H_0 : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$.

Under H_0, our test statistic t has a t-distribution with 12-2=10 df.

```
> qt(df=10,c(.005,.995))
[1] -3.169273  3.169273
```

At level of significance $\alpha = .01$, the test is "Reject H_0 if $|t| > 3.17$." R told us earlier that $t = 4.103$. We reject H_0, and conclude that $\beta \neq 0$.