Question 12.35

a. $\alpha = 0.10$.
b. H_O:"The proportion of habitats used by the bear is the same as the proportions of available habitat" versus H_A:"The proportion of habitats used by the bear is the same as the proportions of available habitat."
c. A goodness-of-fit test is required because a single categorical variable (habitat use) from a single population (this bear) was measured and the proportions are being compared to a theoretical distribution in the null hypothesis.
d. This is an observational study with randomly selected times for observation.
e. The expected number of observations in each habitat is in proportion to the GIS analysis of habitat availability. Thus, the expected number in each habitat is shown in Table 1 (along with the observed frequencies). Note that these percentages did not need to be adjusted to the number of observations because 100 observations were made. The test statistic computed below should reasonably follow a χ^2-distribution because all of the expected values are greater than five.

Table 1. Observed and expected frequencies for habitat use by a bear.

Habitat	Obs Freq	Exp Freq
Lowland Conifer	47	34
Aspen	12	17
Open Areas	10	12
Upland Hardwoods	21	25
Mixed Upland	10	12
Total	100	100

f. The table of observed frequencies is shown in Table 1.
g. The χ^2 test statistic is thus,

$$\begin{aligned}\chi^2 &= \frac{(47-34)^2}{34} + \frac{(12-17)^2}{17} + \frac{(10-12)^2}{12} + \frac{(21-25)^2}{25} + \frac{(10-12)^2}{12} \\ &= 4.971 + 1.471 + 0.333 + 0.640 + 0.333 \\ &= 7.7478\end{aligned}$$

with $5 - 1 = 4$ df.
h. The p-value is $p = 0.1013$.
i. The H_O is not rejected because the $p - value > \alpha$.
j. The bear appears to use the habitats in proportion to the availability of the habitat.

Question 12.36

a. $\alpha = 0.05$.
b. H_0:"The proportion of ginseng plants consumed by deer is 0.33" versus H_A:"The proportion of ginseng plants consumed by deer is NOT 0.33".
c. A goodness-of-fit test is required because a categorical variable (consumed or not) from a single population was recorded and the frequency of responses is being compared to a theoretical distribution in the null hypothesis.
d. An observational study with randomly selected plants was used.
e. The expected number of consumed plants is $73 * 0.33$=24.09. The expected number of plants not consumed by deer is $73 * (1 - 0.33)$=48.91. the test statistic computed below should reasonably follow a χ^2-distribution because both of these expected values is greater than five.
f. The table of observed frequencies is in Table 2 (along with the expected frequencies from the previous step).

Table 2. Observed and expected frequencies for browsing by deer on ginseng plants.

Consumed	Obs Freq	Exp Freq
Yes	33	24.09
No	40	48.91
Total	73	73

g. The χ^2 test statistic is 4.919 with 1 df (Table 3).

Table 3. Results from the goodness-of-fit test for testing that more than one-third of ginseng plants were browsed by deer.

```
X-squared = 4.919, df = 1, p-value = 0.02657
```

h. The p-value is $p = 0.0266$ (Table 3).
i. The H_O is rejected because the $p - value < \alpha$.
j. It appears that more than 33% of the ginseng plants at this site and year were consumed by deer.
k. One is 95% confident that the proportion of ginseng plants consumed by deer is between 0.343 and 0.566

Question 12.37

a. $\alpha = 0.05$.
b. H_O:"The proportion of road rage incidents is the same on each day of the week" versus H_A:"The proportion of road rage incidents is NOT the same on each day of the week".
c. A goodness-of-fit test is required because a categorical response variable with seven levels (days of the week) from a single population was recorded and the frequency of responses is being compared to a theoretical distribution in the null hypothesis.
d. This is an observational study without obvious randomization.
e. The expected number of road rage incidents on each day is $\frac{69}{7} = 9.9$ which are all greater than five. Thus, the test statistic computed below should reasonably follow a χ^2-distribution.
f. The table of observed frequencies is shown in Table 4

Table 4. Observed frequencies of road rage by day of the week.

```
 M Tu  W Th  F Sa Su
 4 11 12 10 18  7  5
```

g. The χ^2 test statistic is 14.388 with 6 df (Table 5).

Table 5. Results from the goodness-of-fit test for testing that the frequencies of road rage differed among days of the week.

```
X-squared = 14.39, df = 6, p-value = 0.02559
```

h. The p-value is $p = 0.0256$ (Table 5).
i. The H_O is rejected because the $p - value < \alpha$.
j. It appears that the proportion of incidences of road rage differs among the days of the week.
k. An examination of the confidence interval results (Table 6) suggests that there are more incidences of road rage on Fridays and possibly fewer on Mondays than would be expected.

Table 6. Confidence intervals (95%) for the proportion of observations of road rage by day of the week.

```
   p.obs p.LCI p.UCI p.exp
M   0.06  0.02  0.14  0.14
Tu  0.16  0.09  0.27  0.14
W   0.18  0.11  0.29  0.14
Th  0.15  0.08  0.25  0.14
F   0.27  0.18  0.39  0.14
Sa  0.10  0.05  0.20  0.14
Su  0.07  0.03  0.16  0.14
```

Appendix – R Commands

```
( distrib(7.7478,distrib="chisq",df=4,lower.tail=FALSE) )
```

```
obs <- c(consumed=33,not.consumed=40)
exp <- c(consumed=0.33,not.consumed=1-0.33)
( gin.chi <- chisq.test(obs,p=exp,correct=FALSE) )
gin.chi$expected
gofCI(gin.chi)
```

```
rr <- read.table("http://www.ncfaculty.net/dogle/Data/R/RoadRage.txt",header=TRUE)
rr$Day1 <- factor(rr$Day,levels=c("M","Tu","W","Th","F","Sa","Su"))
( obs <- table(rr$Day1) )
exp <- c(69/7,69/7,69/7,69/7,69/7,69/7,69/7)
( rr.chi <- chisq.test(obs,p=exp,correct=FALSE,rescale=TRUE) )
rr.chi$expected
rr.chi$residuals
gofCI(rr.chi)
```

Notes From Professor

-