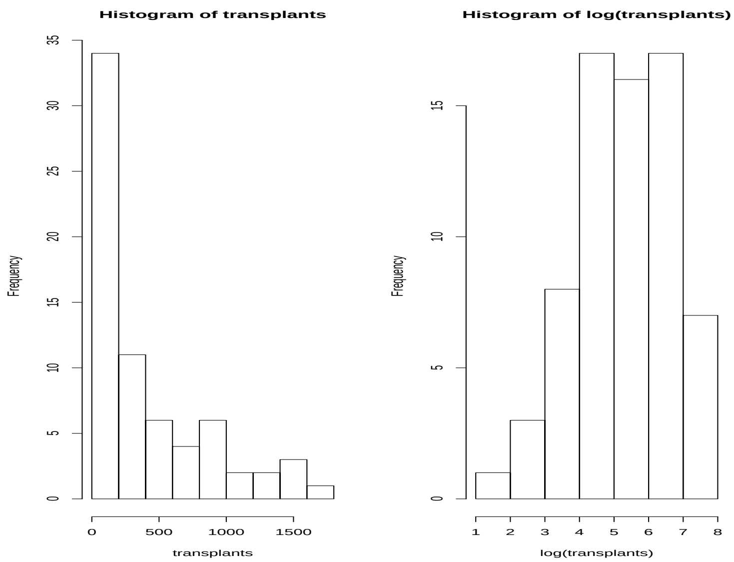
1. Use two data sets called `heart1.dat` and `heart2.dat` from Koehler's Stat 511 page.

   (a) The sample median of transplant survival times is 207 days. The histogram of the survival times for heart transplant patients is skewed to the right. On the other hand, the histogram of the logarithm of the survival times has a more symmetric distribution. However, this does not necessarily mean that it might be described as having a normal distribution.



   We might report the sample standard deviation of the bootstrap sample (67 days) as a standard error for the sample median.

   (b) A 95% (unadjusted) percentile bootstrap confidence interval for the median of F is $(109, 342)$ days.

   (c) The maximum likelihood estimates of the parameters of the Weibull are, shape = 0.86 and scale = 383.3. A parametric bootstrap standard error for the sample median is 50 days and a parametric bootstrap 95% (unadjusted) percentile bootstrap confidence interval for the median of the survival times of heart transplant patients is $(166, 360)$ days.

   Note that the standard error is smaller and the confidence interval is narrower in the parametric bootstrap (compared to the values obtained in 1(b)).

   (d) The density function of a Weibull distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ is $f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)$. The standard deviation of the sample median of a Weibull(0.86,383.3) for $n = 69$ is 44 days.

```
a <- fit1$estimate[1] ; b <- fit1$estimate[2] ; median <- median(transplants)
f <- (a/b)*(median/b)^(a-1)*exp(-(median/b)^a)
1/(2*f*sqrt(69))               # standard deviation of the sample median
```

   (e) 95% percentile confidence limits for the difference in underlying median survival times is $(81, 326)$ days. It is clear that heart transplant patients have a higher survival time.

```
contboot.non <- bootstrap(controlcases,B,"median")
diff <- tranboot.non - contboot.non
kl <- floor((B+1)*.025) ; ku <- B+1-kl
sdiff <- sort(diff)
sdiff[kl] ; sdiff[ku]
```

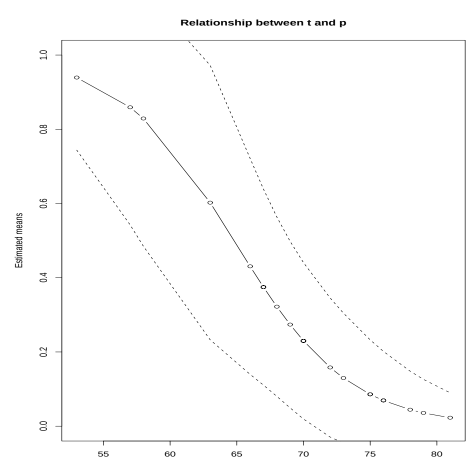
2. Study the solution to Problem 5 of Koehler's Spring 2002 Assignment #9.

3. Use data set of pre-Challenger space shuttle flights.

(a)
```
> summary(shuttle.out)

Call:
glm(formula = indicate ~ temp, family = binomial)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0611  -0.7613  -0.3783   0.4524   2.2175

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  15.0429     7.3719   2.041   0.0413 *
temp         -0.2322     0.1081  -2.147   0.0318 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 28.267  on 22  degrees of freedom
Residual deviance: 20.315  on 21  degrees of freedom
AIC: 24.315

Number of Fisher Scoring iterations: 4
```

There is evidence that the coefficient of the temperature covariate is non zero (p-value=0.03). The test $H_0 : \beta_1 < 0$ has a p-value of 0.015. I would say that this data set suggests that their claim was not correct.

```
shuttle.fits <- predict.glm(shuttle.out,type="response",se.fit=TRUE)
ind <- order(temp)
plot(temp[ind],shuttle.fits$fit[ind],type="b",ylim=c(0,1),
  xlab="Temperature",ylab="Estimated means",main="Relationship between t and p")
lines(temp[ind],shuttle.fits$fit[ind]-2*shuttle.fits$se.fit[ind],lty=2)
lines(temp[ind],shuttle.fits$fit[ind]+2*shuttle.fits$se.fit[ind],lty=2)
```



The temperature 31^{o}F is outside the range of temperature values used to fit the model. However, assuming that the relationship of temperature and O-ring incidents remains the same at lower temperature values, the model suggests that there is a very high probability of having an O-ring incident on a lauch at 31^{o}F.

```
> predict.glm(shuttle.out,data.frame(temp=31),se.fit=TRUE,type="response")
$fit
[1] 0.9996088
$se.fit
[1] 0.001578722
```

4. Use data set gathered in a project aimed at reducing jams on a large collating machine.

(a) It appears that there are statistically detectable Air Pressure and Bar Tightness effects in these data since the coefficients for the levels of these two variables are significant. If one wants small number of jams, one wants level 2 of Air Pressure and level 2 of Bar Tightness.

```
> summary(collator.out)

Call:
glm(formula = y ~ AA + BB, family = poisson, offset = log(k))

Deviance Residuals:
      1        2        3        4        5        6
-0.5040   0.1693   0.3442   0.7453  -0.3004  -0.5532

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -3.2159     0.1068 -30.099  < 2e-16 ***
AA1           0.2420     0.1313   1.844  0.06520 .
AA2          -0.4856     0.1472  -3.299  0.00097 ***
BB1           0.6781     0.1045   6.491 8.54e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 59.201  on 5  degrees of freedom
Residual deviance:  1.353  on 2  degrees of freedom
AIC: 37.188

Number of Fisher Scoring iterations: 3
```

(b)
```
mu <- collator.out$coefficients[1]
alpha1 <- collator.out$coefficients[2] ; alpha2 <- collator.out$coefficients[3]
alpha3 <- -(alpha1+alpha2)
beta1 <- collator.out$coefficients[4] ; beta2 <- -beta1
> exp(c(mu+alpha1+beta1,mu+alpha2+beta1,mu+alpha3+beta1,
        mu+alpha1+beta2,mu+alpha2+beta2,mu+alpha3+beta2))
 0.10069307  0.04863886  0.10084992  0.02593996  0.01253006  0.02598037
```

(c) `collator.fits$fit` $= k*$(answer in 4(b)).

```
> collator.fits$fit
[1] 29.704457 20.233767 31.061776 12.295543  6.766233 12.938224
```

(d)
```
> lcollator.fits$fit
       1        2        3        4        5        6
3.391297 3.007353 3.435978 2.509237 1.911944 2.560186
```