Nathan Foster
Math 728
Homework 1
02/07/2013

Exercise 4.1.2. *The weights of 26 professional baseball pitchers are given. Suppose we assume that the weight of a professional baseball pitcher is normally distributed with mean μ and variance σ^2.*

a) Obtain a frequency distribution and a histogram or a stem-leaf plot of the data. Use 5-pound intervals. Based on this plot is a normal probability model credible?

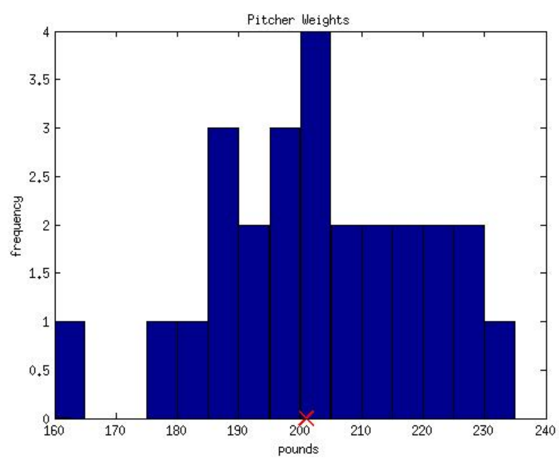
b) Obtain the maximum likelihood estimates of μ, σ^2, σ, and μ/σ. Locate your estimate of μ on your plot in part (a).

c) Using the binomial model, obtain the maximum likelihood estimate of the proportion p of professional baseball pitchers who weigh over 215 pounds.

d) Determine the mle of p assuming that the weight of a professional baseball player follows the normal probability model $N(\mu, \sigma^2)$ with μ and σ unknown.

Solution 4.1.2.

a) Below is the histogram of the data.



The sample size is too small to know if the normal probability model is credible.

b) Below are the maximum likelihood estimates.

$$\hat{\mu} = \bar{X} = 201$$
$$\hat{\sigma}^2 = \frac{n-1}{n} S^2 = 293.9$$
$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = 17.14$$
$$\frac{\hat{\mu}}{\hat{\sigma}} = 11.73$$

Note that $\hat{\mu}$ is plotted on the histogram with a red cross.

c) By example 4.1.2, the maximum likelihood estimate is the relative frequency of the pitchers that weigh over 215 pounds, i.e.

$$\hat{p} = \frac{7}{26}.$$

d) We want to estimate $p = P(X > 215)$ where $X \sim N(\mu, \sigma^2)$. Using the maximum likelihood estimates of μ and σ from part (b) we get

$$\begin{aligned} \hat{p} &= P(\hat{X} > 215) \\ &= 1 - P(\hat{X} \leq 215) \\ &= 1 - F_{\hat{X}}(215) \\ &\approx 0.207 \end{aligned}$$

where $\hat{X} \sim N(\hat{\mu}, \hat{\sigma}^2)$.

$\square$

Exercise 4.1.6. *Show that the estimate of the pmf in expression (4.1.9) is an unbiased estimate. Find the variance of the estimator also.*

Solution 4.1.6. We want to show $E[\hat{p}(a_j)] = p(a_j)$. Let $X_1, \ldots, X_n$ be a discrete random sample, then

$$\begin{aligned}
E[\hat{p}(a_j)] &= E\left[\frac{1}{n}\sum_{i=1}^{n} I_j(X_i)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} E[I_j(X_i)] \\
&= \frac{1}{n}\sum_{i=1}^{n} p(a_j) \\
&= \frac{1}{n} np(a_j) \\
&= p(a_j).
\end{aligned}$$

Also,

$$\begin{aligned}
\mathrm{Var}(\hat{p}(a_j)) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} I_j(X_i)\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(I_j(X_i)) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} p(a_j)(1-p(a_j)) \\
&= \frac{p(a_j)(1-p(a_j))}{n}.
\end{aligned}$$

□

Exercise 4.4.5. *Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample of size 4 from the distribution having pdf $f(x) = e^{-x}, 0 < x < \infty$, zero elsewhere. Find $P(Y_4 \geq 3)$.*

Solution 4.4.5. The pdf of Y_4 can be calculated using equation (4.4.2). So

$$\begin{aligned}
g_4(y_4) &= 4\left(\int_0^{y_4} e^{-x}dx\right)^3 e^{-y_4} \\
&= 4\left(1-e^{-y_4}\right)^3 e^{-y_4}, \quad y_4 > 0.
\end{aligned}$$

Using the change of variables $u = 1 - e^{-y}$, the cdf can be calculated as

$$\begin{aligned} F_{Y_4}(t) &= 4\int_0^t (1-e^{-y})^3 e^{-y} dy \\ &= 4\int_0^{1-e^{-t}} u^3 du \\ &= (1-e^{-t})^4, \quad t > 0. \end{aligned}$$

Hence,

$$P(Y_4 \geq 3) = 1 - F_{Y_4}(3) = 1 - (1-e^{-3})^4 \approx 0.185$$

□

Exercise 4.4.9. *Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample of size n from a distribution with pdf $f(x) = 1, 0 < x < 1$, zero elsewhere. Show that the kth order statistic Y_k has a beta pdf with parameters $\alpha = k$ and $\beta = n - k + 1$.*

Solution 4.4.9. Using result (4.4.2),

$$\begin{aligned} g_k(y_k) &= \frac{n!}{(k-1)!(n-k)!}\left[\int_0^{y_k} 1dx\right]^{k-1}\left[1 - \int_0^{y_k} 1dx\right]^{n-k} \qquad (1) \\ &= \frac{n!}{(k-1)!(n-k)!} y_k^{k-1}(1-y_k)^{n-k} \\ &= \frac{\Gamma(\alpha+\beta-1)}{\Gamma(\alpha-1)\Gamma(\beta-1)} y_k^{\alpha-1}(1-y_k)^{\beta-1}, \quad 0 < y_k < 1. \end{aligned}$$

□

Exercise 5.1.2. *Let the random variable Y_n have a distribution that is $b(n, p)$.*

a) Prove that Y_n/n converges in probability to p. This result is one form of the week law of large numbers.

b) Prove that $1 - Y_n/n$ converges in probability to $1 - p$.

c) Prove that $(Y_n/n)(1 - Y_n/n)$ converges in probability to $p(1 - p)$.

Solution 5.1.2.

a) Let $\epsilon > 0$. Then by Chebyshev's inequality,

$$\begin{aligned} P(|Y_n/n - p| \geq \epsilon) &\leq \frac{\text{Var}(Y_n/n)}{\epsilon^2} \\ &= \frac{np(1-p)}{n^2\epsilon^2} \\ &\to 0 \end{aligned}$$

as $n \to \infty$. Hence, $Y_n/n \xrightarrow{P} p$.

b) Since $1 \xrightarrow{P} 1$ and $Y_n/n \xrightarrow{P} p$, it follows from Theorem 5.1.2 and 5.1.3 that $1 - Y_n/n \xrightarrow{P} 1 - p$.

c) Similarly, it follows from Theorem 5.1.5 that $(Y_n/n)(1 - Y_n/n) \xrightarrow{P} p(1-p)$

□

Exercise 5.1.3. *Let W_n denote a random variable with mean μ and variance b/n^p, where $p > 0$, μ, and b are constants. Prove that W_n converges in probability to μ.*

Solution 5.1.3. Let $\epsilon > 0$. Then by Chebyshev's inequality,

$$\begin{aligned} P(|W_n - \mu| \geq \epsilon) &\leq \frac{\text{Var}(W_n)}{\epsilon^2} \\ &= \frac{b}{n^p\epsilon^2} \\ &\to 0 \end{aligned}$$

as $n \to \infty$. Hence, $W_n \xrightarrow{P} \mu$. □

Exercise 5.2.1. *Let $\bar{X}_n$ denote the mean of a random sample of size n from a distribution that is $N(\mu, \sigma^2)$. Find the limiting distribution of $\bar{X}_n$.*

Solution 5.2.1. By Corollary 3.4.1, $\bar{X}_n \sim N(\mu, \sigma^2/n)$. So

$$\begin{aligned} M_{X_n}(t) &= e^{\mu t + \frac{\sigma^2 t^2}{2n}} \\ &= e^{\mu t} e^{\frac{\sigma^2 t^2}{2n}} \\ &\to e^{\mu t} \end{aligned}$$

as $n \to \infty$. Note that $e^{\mu t}$ is the mgf of the degenerative distribution located at μ. □