prepared by Professor Jenny Baglivo

4 MATH4427 Notebook 4

This notebook is concerned with the use of order statistics to answer questions about quantiles of continuous distributions. The notes include material from Chapter 3 (joint distributions), Chapter 9 (probability plots) and Chapter 10 (summarizing data) of the Rice textbook.

4.1 Kth Order Statistics and Their Distributions

4.1.1 Definitions

Let X_1, X_2, $\ldots$, X_n be a random sample from the continuous distribution

whose PDF is $f(x)$, and whose CDF is $F(x) = P(X \leq x)$, for all real numbers x,

and let k be an integer between 1 and n, $k \in \{1, 2, \ldots, n\}$.

1. *Kth Order Statistic:* The k^{th} *order statistic*, $X_{(k)}$, is the k^{th} observation in order:

$$X_{(k)} \text{ is the } k^{\text{th}} \text{ smallest of } X_1,\ X_2,\ \ldots,\ X_n.$$

2. *Sample Maximum/Minimum:* The largest observation, $X_{(n)}$, is called the *sample maximum* and the smallest observation, $X_{(1)}$, is called the *sample minimum*.
3. *Sample Median:* The *sample median* is the middle order statistic when n is odd, and the average of the two middle order statistics when n is even:

$$\textit{Sample Median} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{when } n \text{ is odd} \\ \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right) & \text{when } n \text{ is even} \end{cases}$$

For example, if the following numbers were observed (and ordered)

13.8, 17.0, 21.0, 24.5, 31.8, 39.5, 45.2, 47.6,

the observed sample minimum is 13.8,

the observed sample maximum is 47.6, and

the observed sample median is ____.

4.1.2 Distribution of the Kth Order Statistic

Let $X_{(k)}$ be the k^{th} order statistic of a random sample of size n from a continuous distribution with PDF $f(x)$ and CDF $F(x)$, where $k \in \{1, 2, \ldots, n\}$. Then

1. *CDF:* The cumulative distribution function of $X_{(k)}$ has the following form:

$$F_{(k)}(x) = \sum_{j=k}^{n} \binom{n}{j} (F(x))^j (1 - F(x))^{n-j}, \quad \text{for all real numbers } x.$$

2. *PDF:* The probability density function of $X_{(k)}$ has the following form:

$$f_{(k)}(x) = \frac{d}{dx} F_{(k)}(x) = \binom{n}{k-1, 1, n-k} (F(x))^{k-1} f(x) (1 - F(x))^{n-k}$$

   (after simplification), whenever the derivative exists.

To demonstrate that the formula for $F_{(k)}(x) = P(X_{(k)} \leq x)$ is correct, first note that the event that "$X_{(k)} \leq x$" is equivalent to the event that "k or more of the X_i's are $\leq x$".

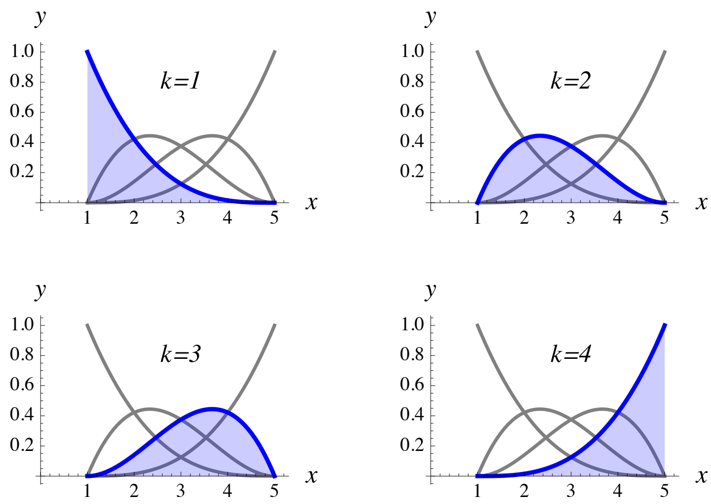
Now (complete the demonstration),

Exercise. Let X be the continuous uniform random variable on the interval $[a, b]$.

(a) Let $X_{(k)}$ be the k^{th} order statistic of a random sample of size n from the X distribution. Find simplified expressions for

$$f_{(k)}(x) \text{ and } F_{(k)}(x) \text{ when } x \in [a, b].$$

(b) Let $[a,b]=[1,5]$ and assume that $n=4$.



The plots above show the density functions for the 4 order statistics. Use your work from part (a) to find

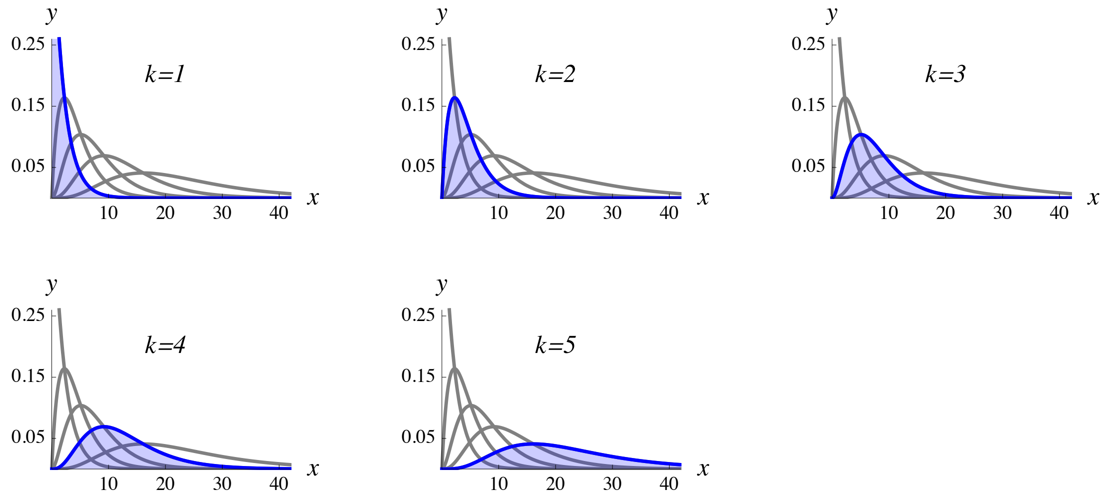
$$P(X_{(2)} \leq 2), \quad P(X_{(2)} \leq 3), \quad P(2 \leq X_{(2)} \leq 3).$$

Exercise. Let X be an exponential random variable with parameter λ.

(a) Let $X_{(k)}$ be the k^{th} order statistic of a random sample of size n from the X distribution. Find simplified expressions for

$$f_{(k)}(x) \text{ and } F_{(k)}(x) \text{ when } x \in (0, \infty).$$

(b) Let $\lambda = \frac{1}{10}$ and assume that $n = 5$.



The plots above show the density functions for the 5 order statistics.

Use your work from part (a) to find

$$P(X_{(4)} \le 10), \quad P(X_{(4)} \le 20), \quad P(10 \le X_{(4)} \le 20).$$

Exercise: Sample Maximum/Minimum. The sample maximum and sample minimum of a random sample of size n from the X distribution are used in many computations.

(a) Simplify the general formulas for $F_{(n)}(x)$ and $f_{(n)}(x)$ as much as possible.

(b) Simplify the general formulas for $F_{(1)}(x)$ and $f_{(1)}(x)$ as much as possible.

Exercise: Quantiles of Sample Maximum/Minimum for uniform distributions. Let X be a uniform random variable on the interval $[a, b]$, and consider the sample maximum and sample minimum of a random sample of size n from the X distribution.

(a) Find a general formula for the p^{th} quantile of the $X_{(n)}$ distribution.

(b) Let $[a, b] = [1, 5]$ and assume $n = 4$. Use your answer to part (a) to find the median of the distribution of the sample maximum.

(c) Find a general formula for the p^{th} quantile of the $X_{(1)}$ distribution.

(d) Let $[a, b] = [1, 5]$ and assume $n = 4$. Use your answer to part (c) to find the median of the distribution of the sample minimum.

4.1.3 Approximate Mean and Variance of Order Statistic Distributions

Summary measures of order statistic distributions are related to the quantiles of X.

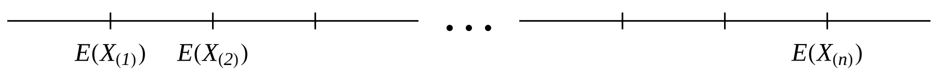
Specifically, let X be a continuous random variable with PDF $f(x)$, $X_{(k)}$ be the k^{th} order statistic of a random sample of size n from the X distribution, $p = k/(n+1)$ and θ be the p^{th} quantile of the distribution of X.

> ***Summary Measures Theorem.*** Under the conditions above, if $f(\theta) \neq 0$, then
>
> $$E(X_{(k)}) \approx \theta \quad \text{and} \quad Var(X_{(k)}) \approx \frac{p(1-p)}{(n+2)(f(\theta))^2}.$$
>
> In these formulas, $f(x)$ is the PDF of X and θ is the p^{th} quantile of the X distribution, where $p = k/(n+1)$.

Notes:

1. The theorem tells us that the $(n+1)$ intervals

   $$(-\infty, E(X_{(1)})), \quad (E(X_{(1)}, E(X_{(2)})), \quad \ldots, \quad (E(X_{(n-1)}, X_{(n)})), \quad (E(X_{(n)}), \infty)$$

   are approximately equally likely. That is,

   $$P(X \in \text{The } i^{\text{th}} \text{ Interval}) \approx \frac{1}{n+1}, \quad \text{for } i = 1, 2, \ldots, n+1.$$

2. The formulas given in the theorem are <u>*exact*</u> for uniform distributions.

To illustrate the theorem for uniform distributions, let $[a, b] = [1, 5]$ and assume that $n = 4$. Summary measures for the 4 order statistics are given below:

	$E(X_{(k)})$	$Var(X_{(k)})$
$k = 1$	9/5	32/75
$k = 2$	13/5	16/25
$k = 3$	17/5	16/25
$k = 4$	21/5	32/75

4.1.4 Graphical Analysis: Probability Plots

Let X be a continuous random variable, and $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the observed values of the order statistics of a random sample of size n from the X distribution.

A *probability plot* is a plot of pairs of the form:

$$\left(\left(\frac{k}{n+1}\right)^{\text{st}} \text{ model quantile, } x_{(k)}\right), \quad k = 1, 2, \ldots, n .$$

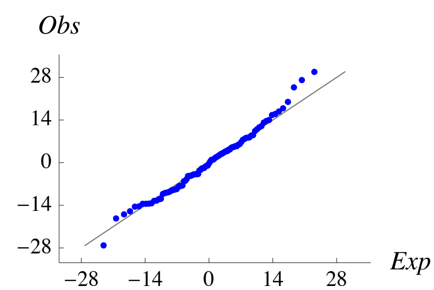
The theorem in the last section tells us that the ordered pairs in a probability plot should lie roughly on the line $y = x$.

For example, I used the computer to generate a pseudo-random sample of size 95 from the normal distribution with mean 0 and standard deviation 10.

1. *Probability Plot:* The plot on the right is a probability plot of pairs of the form

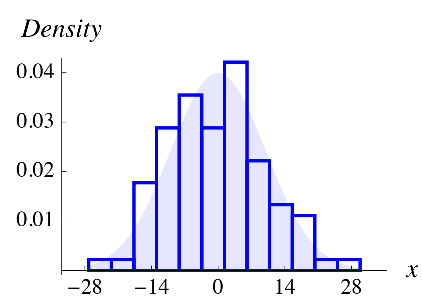
$$\left(\left(\frac{k}{96}\right)^{\text{th}} \text{ model quantile, } x_{(k)}\right)$$

for $k = 1, 2, \ldots, 95$. In the plot, the Observed order statistic (vertical axis) is plotted against its approximate Expected value (horizontal axis).



2. *Comparison Plot:* The plot on the right shows an empirical histogram of the same sample, superimposed on the density curve for a normal distribution with mean 0 and standard deviation 10.

   Twelve subintervals of equal length were used to construct the empirical histogram.



Note: If n is large, then both plots give good graphical comparisons of model and data.

But, if n is small to moderate, then the probability plot may be a better way to compare model and data since the shape of the empirical histogram may be very different from the shape of the density function of the continuous model.

4.2 Estimation and Hypothesis Testing Methods

4.2.1 Large Sample Theory: Sample Median

The following theorem tells us that the sampling distribution of the sample median is approximately normal when n is a large, odd integer.

> ***Sampling Distribution Theorem.*** Let X be a continuous random variable with density function $f(x)$, and with median θ. Further, let n be an odd integer and
>
> $$\widehat{\theta} = X_{\left(\frac{n+1}{2}\right)} \text{ be the sample median of a random sample of size } n.$$
>
> If $f(\theta) \neq 0$ and n is large, then the distribution of
>
> $$\widehat{\theta} \text{ is approximately normal with mean } \theta \text{ and variance } \frac{1}{4n(f(\theta))^2}.$$

4.2.2 Approximate Confidence Interval Procedure for the Median

Under the conditions of the theorem above, an approximate $100(1-\alpha)\%$ confidence interval for the median of the X distribution has the following form:

$$\widehat{\theta} \;\pm\; z(\alpha/2)\sqrt{\frac{1}{4n(\widehat{f(\theta)})^2}} \qquad \text{where } \widehat{\theta} = X_{\left(\frac{n+1}{2}\right)},$$

and $z(\alpha/2)$ is the $100(1-\alpha/2)\%$ point of the standard normal distribution. In this formula, $\widehat{f(\theta)}$ is the estimate of $f(\theta)$ obtained by substituting the sample median for θ.

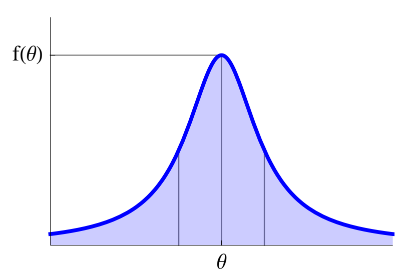
Exercise. Let X be a Cauchy random variable with center θ and spread 1.

The density function of X is

$$f(x) = \frac{1}{\pi(1+(x-\theta)^2)}$$

for all real numbers x.

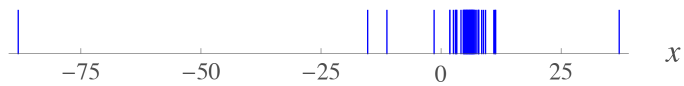
The median of the Cauchy distribution is θ; the mean is indeterminate.



(a) Evaluate the variance formula $\frac{1}{4n(f(\theta))^2}$.

(*b*) Assume the following data are the values of a random sample of size 75 from the Cauchy distribution with center θ and spread 1.

−88.043	−15.284	−11.278	−1.451	1.830	1.838	2.571	2.987	3.056	3.188
3.275	4.145	4.610	4.681	4.705	4.839	4.909	4.922	4.978	5.048
5.165	5.193	5.302	5.308	5.380	5.401	5.412	5.443	5.457	5.636
5.652	5.792	5.794	5.822	5.826	5.935	5.944	5.985	5.991	6.059
6.170	6.185	6.239	6.255	6.301	6.350	6.351	6.411	6.448	6.449
6.485	6.515	6.521	6.556	6.571	6.600	6.784	6.846	6.866	6.869
6.885	7.215	7.302	7.693	7.776	7.833	7.844	8.455	8.828	9.258
10.984	11.232	11.270	11.365	37.091					



Use the normal approximation to the distribution of the sample median to construct an approximate 90% confidence interval for θ.

4.2.3 Exact Confidence Interval Procedures for Quantiles

Let X be a continuous random variable, and let θ be the p^{th} quantile of the X distribution, for some proportion $p \in (0,1)$.

Let $X_{(k)}$ be the k^{th} order statistic of a random sample of size n from the X distribution. Then

1. *Intervals:* The n order statistics divide the real line into $(n+1)$ intervals:

$$(-\infty, X_{(1)}),\quad (X_{(1)}, X_{(2)}),\quad \ldots,\quad (X_{(n-1)}, X_{(n)}),\quad (X_{(n)}, \infty)$$

   (ignoring the endpoints).

2. *Binomial Probabilities:* The probability that θ lies in a given interval follows a binomial distribution with parameters n and p. Specifically,

   (a) *First Interval:* The event "$\theta \in (-\infty, X_{(1)})$" is equivalent to the event that all X_i's are greater than θ. Thus,

$$P(\theta \in (-\infty, X_{(1)})) = (1-p)^n.$$

   (b) *Middle Intervals:* The event "$\theta \in (X_{(k)}, X_{(k+1)})$" is equivalent to the event that exactly k X_i's are less than θ. Thus,

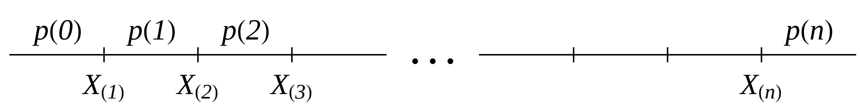
$$P(\theta \in (X_{(k)}, X_{(k+1)})) = \binom{n}{k} p^k (1-p)^{n-k}.$$

   (c) *Last Interval:* The event "$\theta \in (X_{(n)}, \infty)$" is equivalent to the event that all X_i's are less than θ. Thus,

$$P(\theta \in (X_{(n)}, \infty)) = p^n.$$

Let $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \ldots, n$, be the binomial probabilities.

Then the following graphic illustrates the probabilities associated with each subinterval.



Further, the facts above can be used to construct a confidence interval procedure for quantiles.

Quantile Confidence Interval Theorem. Under the conditions above, if indices k_1 and k_2 are chosen so that

$$\begin{aligned} P(\theta < X_{(k_1)}) &= \textstyle\sum_{j=0}^{k_1-1} \binom{n}{j} p^j (1-p)^{n-j} = \alpha/2 \\ P(X_{(k_1)} < \theta < X_{(k_2)}) &= \textstyle\sum_{j=k_1}^{k_2-1} \binom{n}{j} p^j (1-p)^{n-j} = 1-\alpha \\ P(\theta > X_{(k_2)}) &= \textstyle\sum_{j=k_2}^{n} \binom{n}{j} p^j (1-p)^{n-j} = \alpha/2 \end{aligned}$$

then the interval $\left[X_{(k_1)}, X_{(k_2)}\right]$ is a $100(1-\alpha)\%$ confidence interval for θ.

Note that, in practice, k_1 and k_2 are chosen to make the sums in the theorem as close as possible to the values shown on the right.

Exercise (Source: Sheaffer et al, 1996). The following table shows the total yearly rainfall (in inches) for Los Angeles in the 10-year period from the beginning of 1983 to the end of 1992.

	1983	*1984*	*1985*	*1986*	*1987*	*1988*	*1989*	*1990*	*1991*	*1992*
Rainfall	34.04	8.90	8.92	18.00	9.11	11.57	4.56	6.49	15.07	22.56

Assume these data are the values of a random sample from a continuous distribution.

(a) Construct a 90% (or as close as possible) confidence interval for the median rainfall. State the exact confidence level.

$p(0) = 0.001$

$x_{(1)} =$ ________

$p(1) = 0.010$

$x_{(2)} =$ ________

$p(2) = 0.044$

$x_{(3)} =$ ________

$p(3) = 0.117$

$x_{(4)} =$ ________

$p(4) = 0.205$

$x_{(5)} =$ ________

$p(5) = 0.246$

$x_{(6)} =$ ________

$p(6) = 0.205$

$x_{(7)} =$ ________

$p(7) = 0.117$

$x_{(8)} =$ ________

$p(8) = 0.044$

$x_{(9)} =$ ________

$p(9) = 0.010$

$x_{(10)} =$ ________

$p(10) = 0.001$

(b) Construct a 90% (or as close as possible) confidence interval for the 40^{th} percentile of the rainfalls distribution. State the exact confidence level.

$p(0) = 0.0060$	
	$x_{(1)} =$ _______
$p(1) = 0.0403$	
	$x_{(2)} =$ _______
$p(2) = 0.1209$	
	$x_{(3)} =$ _______
$p(3) = 0.2150$	
	$x_{(4)} =$ _______
$p(4) = 0.2508$	
	$x_{(5)} =$ _______
$p(5) = 0.2007$	
	$x_{(6)} =$ _______
$p(6) = 0.1115$	
	$x_{(7)} =$ _______
$p(7) = 0.0425$	
	$x_{(8)} =$ _______
$p(8) = 0.0106$	
	$x_{(9)} =$ _______
$p(9) = 0.0016$	
	$x_{(10)} =$ _______
$p(10) = 0.0001$	

4.2.4 Procedures for Endpoint Parameters

Let X be a continuous random variable.

1. *Upper Endpoint/Sample Maximum:* If the range of X has upper endpoint θ, then the sample maximum can be used in statistical procedures concerning θ.

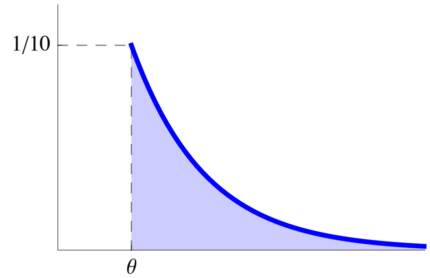
2. *Lower Endpoint/Sample Minimum:* If the range of X has lower endpoint θ, then the sample minimum can be used in statistical procedures concerning θ.

The following multipart exercise illustrates the use of the sample minimum.

Exercise. Let X be the continuous random variable with density function

$$f(x) = \frac{1}{10}\, e^{-\left(\frac{x-\theta}{10}\right)} \quad \text{when } x > \theta, \text{ and } 0 \text{ otherwise.}$$

Note that this X has a *shifted exponential distribution* with shift parameter θ and scale parameter $\lambda = \frac{1}{10}$.



Let $X_{(1)}$ be the sample minimum of a random sample of size n from the X distribution.

(*a*) Completely specify the CDF of $X_{(1)}$: $F_{(1)}(x) = P(X_{(1)} \leq x)$.

(*b*) Find a general formula for the p^{th} quantile of the $X_{(1)}$ distribution.

(c) Let x_p and x_{1-p} be the p^{th} and $(1-p)^{\text{th}}$ quantiles of the $X_{(1)}$ distribution.

Use the fact that $1 - 2p = P(x_p \leq X_{(1)} \leq x_{1-p})$ to fill-in the following blanks:

$$P\Big(\underline{\qquad} \leq \theta \leq \underline{\qquad} \Big) = 1 - 2p.$$

(d) Assume the following data are the values of a random sample from the X distribution:

$$25.8277,\ 28.5495,\ 30.1759,\ 35.8057,\ 36.6871,\ 57.257$$

Use the result of part (c) to construct a 90% confidence interval for θ.

(e) Find the value of c so that the test with decision rule

$$\text{Reject } \theta = 15 \text{ in favor of } \theta < 15 \text{ when } X_{(1)} \leq c$$

is a 5% lower tail test of the null hypothesis $\theta = 15$.

(f) Assume the following data are the values of a random sample from the X distribution:

$$15.2564,\ 15.9165,\ 17.9189,\ 20.2458,\ 23.221,\ 25.3941,\ 27.2703,\ 29.8149$$

Would you accept or reject $\theta = 15$ in this case? What is the observed significance level?

4.3 Sample Quantiles

4.3.1 Definitions

Let X be a continuous random variable, let θ_p be the p^{th} quantile of the X distribution, and let $X_{(k)}$ be the k^{th} order statistic of a random sample of size n , for $k = 1, 2, \ldots, n$.

1. p^{th} *Sample Quantile:* For $p \in \left[\frac{1}{n+1}, \frac{n}{n+1}\right]$, the p^{th} sample quantile is defined as follows:

   (a) if $p = \frac{k}{n+1}$ for some k, then $\widehat{\theta_p} = X_{(k)}$;

   (b) if $p \in \left(\frac{k}{n+1}, \frac{k+1}{n+1}\right)$ for some k, then $\widehat{\theta_p} = X_{(k)} + ((n+1)p - k)\left(X_{(k+1)} - X_{(k)}\right)$.

   With this definition, the point $(\widehat{\theta_p}, p)$ is on the piecewise linear curve connecting the successive points

$$\left(X_{(1)}, \frac{1}{n+1}\right), \left(X_{(2)}, \frac{2}{n+1}\right), \ldots, \left(X_{(n)}, \frac{n}{n+1}\right).$$

2. *Sample Quartiles/Median/IQR:* The *sample quartiles* are the estimates of the 25^{th}, 50^{th}, and 75^{th} percentiles:

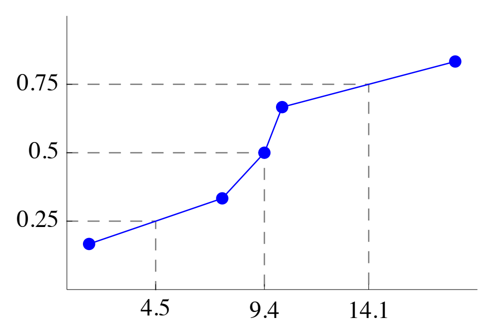
$$q_1 = \widehat{\theta_{0.25}}, \; q_2 = \widehat{\theta_{0.50}}, \; q_3 = \widehat{\theta_{0.75}}.$$

   The *sample median* is q_2, and the *sample interquartile range* is the difference $q_3 - q_1$.

For example, suppose that the 5 numbers

1.5, 7.5, 9.4, 10.2, 18.0

are observed. Then



- the sample median is $x_{(3)} = 9.4$,

- the sample first and third quartiles are

$$q_1 = x_{(1)} + 0.50(x_{(2)} - x_{(1)}) = 4.5 \quad\text{and}\quad q_3 = x_{(4)} + 0.50(x_{(5)} - x_{(4)}) = 14.1,$$

- the sample interquartile range is $q_3 - q_1 = 9.6$.

Exercise (Source: Rice textbook, Chapter 10). As part of a study on the effects of an infectious disease on the lifetimes of guinea pigs, more than 400 animals were infected.

The data below are the lifetimes (in days) of 45 animals given a *low* exposure to the disease:

33	44	56	59	74	77	93	100	102	105	107	107	108	108	109
115	120	122	124	136	139	144	153	159	160	163	163	168	171	172
195	202	215	216	222	230	231	240	245	251	253	254	278	458	555

(a) Find the sample quartiles and sample interquartile range.

(b) Find the sample 90^{th} percentile.

4.3.2 Graphical Analysis: Box Plots

A *box plot* is a graphical display of a data set that shows the sample median, the sample interquartile range, and the presence of possible outliers (numbers that are far from the center). Box plots were introduced by John Tukey in the 1970's.

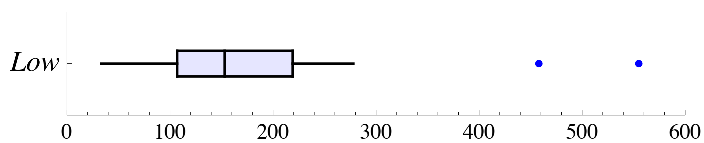
Let q_1, q_2 and q_3 be the sample quartiles. To construct a box plot:

1. *Box:* A box is drawn from q_1 to q_3.
2. *Bar:* A bar is drawn at the sample median, q_2.
3. *Whiskers:* A whisker is drawn from q_3 to the largest observation that is less than or equal to $q_3 + 1.50(q_3 - q_1)$. Another whisker is drawn from q_1 to the smallest observation that is greater than or equal to $q_1 - 1.50(q_3 - q_1)$.
4. *Outliers:* Observations outside the interval

   $$[q_1 - 1.50(q_3 - q_1), q_3 + 1.50(q_3 - q_1)]$$

   are drawn as separate points. These observations are called the *outliers.*

For example, a box plot of the lifetimes data from the last exercise is shown below:



For the lifetimes data, the interval $[q_1 - 1.50(q_3 - q_1), q_3 + 1.50(q_3 - q_1)]$ is

and the outliers are ________________________________ .

Exercise. The following data are birthweights (in ounces) for two groups of children:

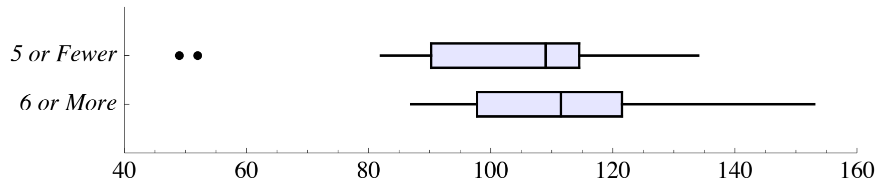
(1) Children whose mothers visited their doctors five or fewer times during pregnancy:

49 52 82 93 96 101 108 110 114 114 114 116 120 134

(2) Children whose mothers visited their doctors six or more times during pregnancy:

87 93 97 98 106 108 110 113 116 119 119 129 131 153

These data are compared using box plots:



Find the sample quartiles (q_1, q_2, q_3), and the interval $[q_1 - 1.50(q_3 - q_1), q_3 + 1.50(q_3 - q_1)]$ for each data set.

Example (Source: Rice textbook, Chapter 10). Consider side-by-side box plots of the following lifetimes (in days) of guinea pigs given *low*, *medium*, and *high* exposure to an infectious disease:

1. *Low Exposure:*

33	44	56	59	74	77	93	100	102	105	107	107	108	108	109
115	120	122	124	136	139	144	153	159	160	163	163	168	171	172
195	202	215	216	222	230	231	240	245	251	253	254	278	458	555

Sample Summaries: $n = 45$, median = 153, iqr = 112

2. *Medium Exposure:*

10	45	53	56	56	58	66	67	73	81	81	81	82	83	88
91	91	92	92	97	99	99	102	102	103	104	107	109	118	121
128	138	139	144	156	162	178	179	191	198	214	243	249	380	522

Sample Summaries: $n = 45$, median = 102, iqr = 69

3. *High Exposure:*

15	22	24	32	33	34	38	38	43	44	54	55	59	60	60
60	61	63	65	65	67	68	70	70	76	76	81	83	87	91
96	98	99	109	127	129	131	143	146	175	258	263	341	341	376

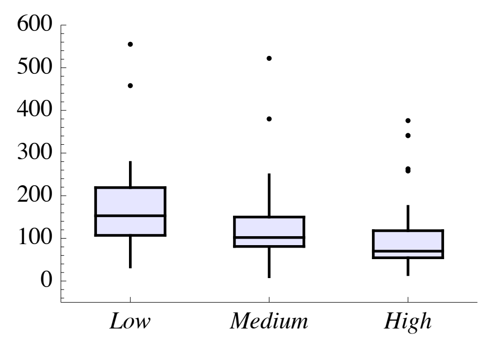
Sample Summaries: $n = 45$, median = 70, iqr = 63.5

The graph suggests a strong relationship between level of exposure and lifetime. Both

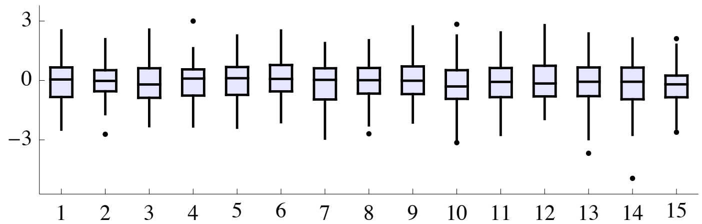
- the sample median lifetime and
- the sample IQR

decrease with increasing exposure.

In addition, as exposure increases, the sample distributions become more skewed.



Exercise. I used the computer to generate 15 random samples, each of size 100, from the standard normal distribution. The samples are displayed as side-by-side box plots.



Notice that the boxes are all reasonably symmetric and approximately centered at 0. Notice also that the boxes have few (if any) outliers.

This problem is concerned with the proportion of observations we might expect to be covered by the box and whiskers, if sampling is done from the standard normal distribution. Specifically, let Z be the standard normal random variable, and

$$w_1 \;=\; z_{0.25} - 1.50\,(z_{0.75} - z_{0.25}) \quad \text{and} \quad w_2 \;=\; z_{0.75} + 1.50\,(z_{0.75} - z_{0.25})\,,$$

where z_p is the p^{th} quantile of the standard normal distribution.

Find $P(w_1 \leq Z \leq w_2)$.