3.2

Assumptions

1. The X_is are a random sample (i.e., they are independent and identically distributed random variables).
2. The measurement scale of the X_is is at least ordinal.

Test Statistic We will use two test statistics in this test. Let T_1 equal the number of observations less than or equal to x^*, and let T_2 equal the number of observations less than x^*. Then $T_1 = T_2$ if none of the numbers in the data exactly equals x^*. Otherwise, T_1 is greater than T_2.

Null Distribution The null distribution of the test statistics T_1 and T_2 is the binomial distribution, with parameters n = sample size, and $p = p^*$ as given in the null hypothesis. The null distribution is given in Table A3 for $n \leq 20$ and selected values of p.

For other values of n and p the normal approximation is used. That is, approximate quantiles x_q for T are given by

$$x_q = n \cdot p + z_q \sqrt{n \cdot p \cdot (1 - p)} \tag{1}$$

where z_q is the qth quantile of a standard normal random variable, given in Table A1.

Hypotheses Let x^* and p^* represent some specified numbers, $0 < p^* < 1$. The hypotheses may take one of the following three forms.

A. (Two-Tailed Test)

$$H_0\text{: The } p^*\text{th population quantile is } x^*$$

[This is equivalent to H_0: $P(X \leq x^*) \geq p^*$, and $P(X < x^*) \leq p^*$, where X has the same distribution as the X_is in the random sample.]

$$H_1\text{: } x^* \text{ is not the } p^*\text{th population quantile}$$

The rejection region corresponds to values of T_2 that are too large [indicating that perhaps $P(X < x^*)$ is greater than p^*] and to values of T_1 that are too small [indicating that perhaps $P(X \leq x^*)$ is less than p^*]. The rejection region is found by entering Table A3 with the sample size n and the hypothesized probability p^*, as in the two-tailed binomial test. Find the number t_1 such that

$$P(Y \leq t_1) = \alpha_1 \tag{2}$$

3.2: The Quantile Test

$\boxed{n \leq 20}$

Ho: The p^{*}th Pop. quantile $\leq x^*$	Ho: The p^* Pop. quant $\geq x^*$	Ho: The p^* Pop. quat $= x^*$
Ha: " " $> x^*$	Ha: " " $< x^*$	Ha: " " $\neq x^*$
$C = \{T_1 : T_1 \leq t_1\}$	$C = \{T_2 : T_2 > t_2\}$	$C = \{T_1 \leq t_1 \text{ or } T_2 > t_2\}$
find t_1 such that $P(y \leq t_1) = \alpha$	find t_2 such that $P(y \leq t_2) = 1-\alpha$	find t_1 & t_2 such $P(y \leq t_1) = \alpha/2$ and $P(y \leq t_2) = 1-\alpha/2$

where

$T_1 = \#$ of $X_i \leq x^*$

$T_2 = $ " " $< x^*$

or using P.Value

$P.V = P(y \leq T_1)$	$P.V = P(y \geq T_2)$	$P.V = 2P(y \leq T_1)$ or $P.V = 2P(y \geq T_2)$ * The smaller

Cont. 3.2: The Quartile Test

$\boxed{n > 20}$

Ho: the p^{α}th Pop. quartile $\leq x^{*}$	Ho: the p^{α}th Pop. quant $\geq x^{\alpha}$	Ho: the p^{α}th Pop. $q_{u} = x^{*}$
Ha: // //// $> x^{\alpha}$	Ha: // // $< x^{\alpha}$	Ha: //// $\neq x^{*}$
$c = \{T_1 : T_1 \leq t_1\}$	$c = \{T_2 : T_2 > t_2\}$	$c = \{T_1 \leq t_1 \text{ or } T_2 > t_2\}$
where $t_1 = n \cdot p^{\alpha} + Z_{\alpha}\sqrt{np^{\alpha}(1-p^{\alpha})}$	where $t_2 = n \cdot p^{\alpha} + Z_{1-\alpha}\sqrt{np^{\alpha}(1-p^{\alpha})}$	where $t_1 = np^{\alpha} + Z_{\alpha/2}\sqrt{np^{\alpha}(1-p^{\alpha})}$ and $t_2 = np^{\alpha} + Z_{1-\alpha/2}\sqrt{np^{\alpha}(1-p^{\alpha})}$

where

$T_1 = \#$ of $x_i \leq x^{*}$

$T_2 =$ //// $< x^{\alpha}$

or using P.Value

$= P(y \leq T_1) = P\left[Z \leq \frac{T_1 - np^{\alpha} + .5}{\sqrt{np^{\alpha}(1-p^{\alpha})}}\right]$ (1)	$= P(y \geq T_2) = 1 - P\left(Z \leq \frac{T_2 - np^{\alpha} - .5}{\sqrt{np^{\alpha}(1-p^{\alpha})}}\right)$ (2)	use (1) & (2) - double them - the smaller

where $1 - \alpha$ is a known *confidence coefficient* and where $X^{(r)}$ and $X^{(s)}$ are order statistics (see Definition 2.1.4) with r and s specified. The values for r and s may be determined prior to drawing the sample in the manner described next, with knowledge of only the sample size n and the desired confidence coefficient. The sample $X_1, \ldots, X_n$ needs only to be random. No restrictions are made on the distribution function of X_i. Thus this statistical method may be applied freely to any random sample from any population.

▶ Confidence Interval for a Quantile

Data The data consist of observations on $X_1, X_2, \ldots, X_n$, which are independent and identically distributed random variables. $X^{(1)} \leq X^{(2)} \leq \cdots \leq X^{(r)} \leq \cdots \leq X^{(s)} \leq \cdots \leq X^{(n)}$ represents the ordered sample, where $1 \leq r < s \leq n$. We wish to find a confidence interval for the (unknown) p^*th quantile, where p^* is some specified number between zero and one.

Assumptions

1. The sample $X_1, X_2, \ldots, X_n$ is a random sample.
2. The measurement scale of the X_is is at least ordinal.

Method A (small samples) For $n \leq 20$ Table A3 may be used to find r and s. Enter Table A3 with the sample size n and the probability $p = p^*$. Read down the column for $p = p^*$ until reaching an entry approximately equal to $\alpha/2$, where $1 - \alpha$ is the approximate confidence coefficient desired. Call that entry α_1, and the corresponding value for y (to the far left of α_1) is $r - 1$. Add 1 to get r. Then continue down the column for $p = p^*$ until reaching an entry approximately equal to $1 - (\alpha/2)$, which we will call $1 - \alpha_2$. The value of y corresponding to the entry $1 - \alpha_2$ is called $s - 1$, and 1 is added to obtain s. Thus we have determined α_1, α_2, r, and s. The exact confidence coefficient is $1 - \alpha_1 - \alpha_2$. The interval estimator is the interval between $X^{(r)}$ and $X^{(s)}$, whose values may be obtained from the data. Then

$n \leq 20$

$$\boxed{P(X^{(r)} \leq x_{p^*} \leq X^{(s)}) \geq 1 - \alpha_1 - \alpha_2} \tag{19}$$

provides the confidence interval. If we assume that the unknown distribution function is continuous, then

$$P(X^{(r)} \leq x_{p^*} \leq X^{(s)}) = 1 - \alpha_1 - \alpha_2 \tag{20}$$

as stated in Equation 18 also.

Method B (large sample approximation) For n greater than 20 the approximation based on the central limit theorem may be used. (See Equation 1.) Compute

n>20

$$\boxed{r^* = np^* + z_{\alpha/2}\sqrt{np^*(1-p^*)}} \tag{21}$$

and

$$\boxed{s^* = np^* + z_{1-\alpha/2}\sqrt{np^*(1-p^*)}} \tag{22}$$

where the quantiles z_p are obtained from Table A1 and where $1-\alpha$ is the desired confidence coefficient. In general, r^* and s^* will not be integers. Let r and s be the integers obtained by rounding r^* and s^* upward to the next higher integers. Then the approximate confidence interval is given by Equation 19, or Equation 20 if the unknown distribution function is continuous.

A one-sided confidence interval may be formed by finding only r or s as described. One-sided confidence intervals are of the form

$$P(X^{(r)} \le x_{p^*}) = 1 - \alpha_1 \tag{23}$$

and

$$P(x_{p^*} \le X^{(s)}) = 1 - \alpha_2 \tag{24}$$

if the distribution function is continuous, or

$$P(X^{(r)} \le x_{p^*}) \ge 1 - \alpha_1 \tag{25}$$

and

$$P(x_{p^*} \le X^{(s)}) \ge 1 - \alpha_2 \tag{26}$$

otherwise.

Computer Assistance *Minitab* (under Median Test) and *StatXact* (under Sign Test) find a confidence interval for the median.

EXAMPLE 3

Sixteen transistors are selected at random from a large batch of transistors and are tested. The number of hours until failure is recorded for each one. We wish to find a confidence interval for the upper quartile, with a confidence coefficient close to 90%. Table A3 is entered with $n = 16$ and $p = 0.75$. Reading down the column for $p = 0.75$ the probability 0.0271 is selected as being close to 0.05. The value of y associated with $\alpha_1 = 0.0271$ is $y = 8$; therefore r equals 9. The probability