

Measuring Pay Equity: Small Sample Size Group Analysis

When sample size is small (below 30/5-rule threshold, [see XXX reference intro](#)), group-level statistical investigations of pay equity should employ methods developed specifically for small samples. One such method that is commonly used is the Test of Median-Split Proportions, which the EEOC refers to as the “Threshold Statistical Test” (see <https://www.eeoc.gov/policy/docs/compensation.html>). The test of median-split proportions for an analysis of gender equity in pay compares the proportion of men and women who are above and below median pay. Gender equity in pay is indicated when men and women are equally represented above the median value, i.e., when the proportions of male and female worker earning above-median pay are equivalent.

An analysis of gender equity in pay using the test of median-split proportions requires the following steps:

Step 1: Identify Employees in Substantially Similar Work Categories

The California Fair Pay Act (FPA) was enacted to ensure that individuals performing "substantially similar work," when viewed as a composite of skill, effort, and responsibility, are compensated equally. It therefore is most appropriate to test for disparities in pay among individuals who are performing substantially similar work ([LINK TO STEP BY STEP](#)).

Step 2: Compute Median Pay

Median pay is the value of the middle case when the cases in the dataset are ordered from the lowest to highest value of the variable measuring pay. Here are the steps to obtain the median pay:

1. Order the cases in the dataset from lowest to highest value of the pay/wage variable
2. Find the mid-point from this range of wages:
 - If there is an odd number of records in the data, the median is the value of the case at the middle of the array.
 - If there is an even number of records in the data, then median is the average of the values for the two cases at the middle of the array.

Step 3: Count

Count the number of men and women above and below the median value. If there is an odd number of records in the dataset, the individual at the median is included in the “median and below” group. Arrange the counts into a 2x2 table.

Figure 1: 2x2 Table of counts of Gender by Median

	Median and Below	Above-Median
Women	<i>Women Below Median</i>	<i>Women Above Median</i>
Men	<i>Men Below Median</i>	<i>Men Above Median</i>

Measuring Pay Equity: Small Sample Size Group Analysis

Step 4: Test for Differences in Representation Above the Median

To test if men and women are equally represented above the median value, statistical methods for testing differences in proportions are used. When the sample size is small, Fisher's Exact Test is generally accepted as the most appropriate method. When the sample size is sufficiently large, the Chi-Square test of Independence is generally preferred. When these tests indicate that there are significant differences in proportions, it may indicate that there is a gender difference pay.

Example illustrating the Test of Median-Split Proportions:

Step 1: We identified 15-people doing substantially similar work and would like to test for gender equity in pay among them. The sample size does not meet the 30/5-rule for regression analysis so we perform a median-split test. Here is the dataset for this hypothetical example:

EMPID	Salary	Gender
1	\$35,900	F
2	\$36,100	F
3	\$36,100	F
4	\$36,100	M
5	\$37,900	F
6	\$38,800	F
7	\$39,300	F
8	\$39,600	F
9	\$39,800	M
10	\$41,000	M
11	\$42,300	M
12	\$43,800	M
13	\$44,100	M
14	\$44,300	F
15	\$44,500	F

Step 2: To compute the median, we first order the cases data from lowest to highest value of Salary and find the mid-point. Since there are an odd number of records, the median is the value of Salary for the 8th case, i.e., it is \$39,600.

Step 3: We count the number of men and women who are above and below the median (\$39,600).

	Median and Below	Above-Median
Women	7	2
Men	1	5

Measuring Pay Equity: Small Sample Size Group Analysis

Step 4: In this analysis, we see that only 22% ($2/9=0.22$) of females are above the median, while 83% ($5/6=0.83$) of males are above the median. Because the counts are small, we analyzed this with the Fisher's Exact Test and obtain $p=0.04$, $SD=2.05$. The compute p -value is less than or equal to 0.05 and the SD is greater than or equal to 1.96, so we can conclude that the difference in proportions women (22%) and men (83%) is statistically significant. This result indicates there may be a significant gender gap in salary among the employees performing substantially similar jobs.

DRAFT