Gender Wage Gap in Armenia
Statistical Analysis Project

Nane Khachatryan, Araksya Gevorgyan

May 10, 2025

1 Introduction

We have heard a lot that in Armenia, women are paid less than men for the same work despite having the same education, skills, and experience. This is not only unfair to women as individuals, but it is a type of discrimination that significantly harms Armenia's economy. But does this gap really exist? If so, what factors affect this? The purpose of this research is to identify if this problem really exists in Armenia, measure how big this wage gap is and identify what drives this inequality. The data is taken from the LFS in Armenia, which is officially accepted by the government. More about the data collection and consistency is provided in the next section.

2 Data Description

The Labour Fource Survey (LFS) is a comprehensive study designed to asses employment, unemployment and the overall labor market in Armenia. The target population sample examined is urban and rural households across Armenia. All people considered are permanent and temporary residents, and those absent from Armenia for less than 12 months. The survey covers the entire territory of the Republic, and the surveyed units are private households (HH) selected randomly, excluding institutional households (e.g,. residential home for the elderly, place of detention, etc.). The survey is conducted through the following tools: HH (main and reserve) list, sampling report (hereinafter referred to as HH list) LFS Questionnaire (Units A- F). The questionary consists of the following sections:

- Section A: COVER SHEET
- Section B: HOUSEHOLD COMPOSITION & CHARACTERISTICS
- Section C: POPULATION MOVEMENT
- Section D: IDENTIFICATION OF EMPLOYMENT
- Section E: MAIN JOB DETAILS

- Section F: SECONDARY JOB
- Section G: WORKING HOURS
- Section H: VOCATIONAL TRAINING
- Section J: JOB SEARCH & AVAILABILITY
- Section K: UNPAID WORK
- Section L: RESPONDENT INFORMATION

As mentioned, each section has multiple questions. You will find the original question sheet attached in the reference.

3 Exploratory Data Analysis

At first, it is important to understand the data we are analyzing. Before the filtration process, different graphs have been made.
In Figure 1, the mean salary of both genders grouped by their jobs with a 95% confidence interval is shown. As you can see, the mean salary of men is higher across all job categories. However, the length of confidence intervals for the mean salary of men is larger, which means we cannot be as certain about the mean salary of men as we can be for that of women.
In Figure 2, the number of men and women working at each job is shown. Among occupations such as agricultural work and service & sales, the number of male and female employees is almost equally distributed. However, men make up the majority of employees in elementary occupations, operators, craft workers and legislators & managers. Conversely, women comprise the majority of employees among clerks, professionals and technicians. Interestingly, the number of women is greater across all categories, but when it comes to the mean salary, the advantage is given to men.
In Figure 3, the wage gap across all regions of Armenia is shown. The wage gap is present in all regions and fluctuates roughly between 50000 and 25000 AMD, with the smallest gap present in Ararat and the largest in Armavir.

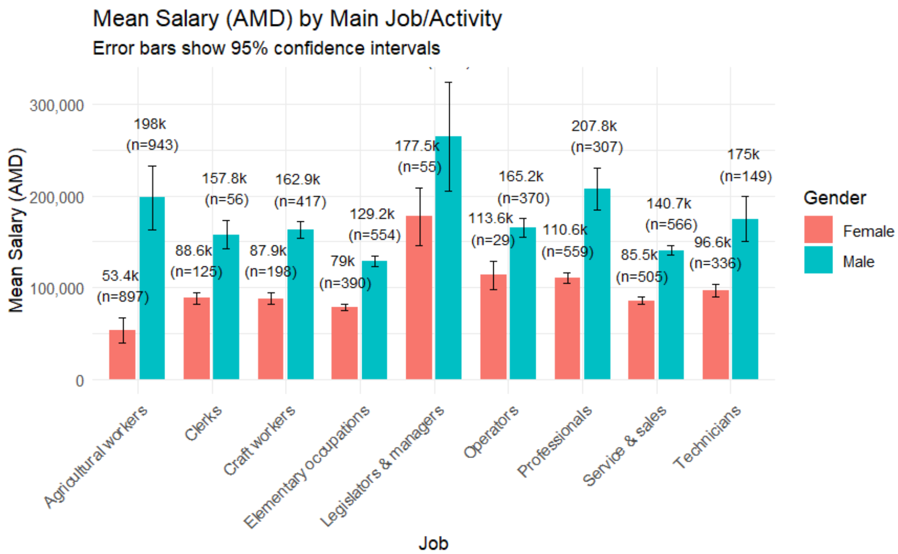


Figure 1:

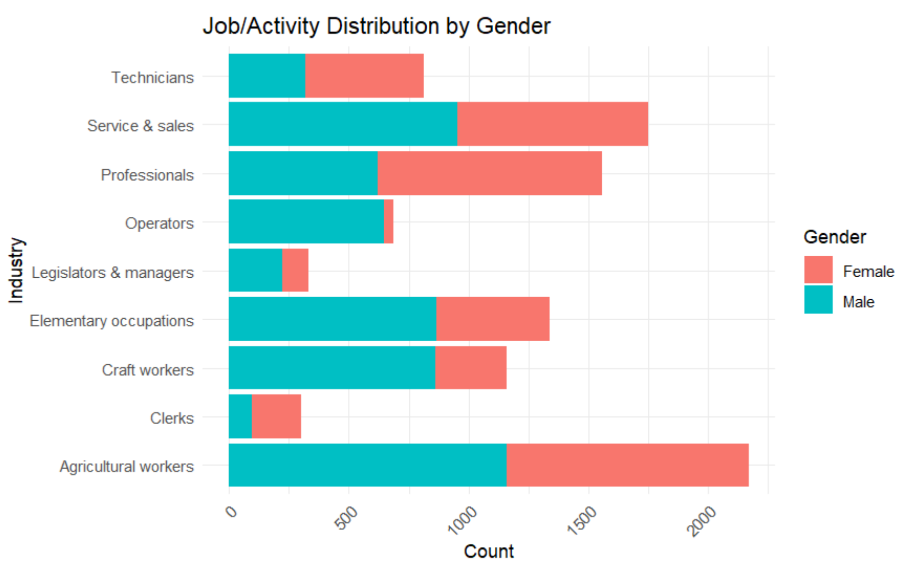


Figure 2:

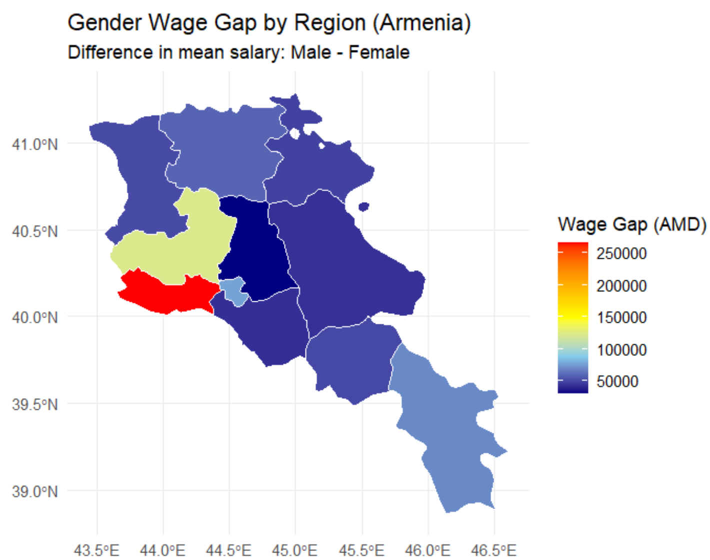


Figure 3:

4 Methodology

Getting back to the primary question raised in the research, the following statistical models are used[1]:

- Two-sample z-test and t-test: Big sample Hypothesis testing for the difference of sample mean
- Two-sample z-test and t-test[2]: Small sample hypothesis testing for the difference of sample mean
- Chi-square test of independence
- Linear regression

Hypothesis tests for different variables:

- Do males earn more than females when the moths do the same job?
- Do males earn more than females when both live in rural/urban areas?
- Do males earn more than females when both have the same level of education?

[1] The significance level for all models is $\alpha = 0.01$

[2] Agresti, A. (2019). An Introduction to categorical data analysis (3rd edn.). Wiley. https://mregresion.wordpress.com/wp-content/uploads/2012/08/agresti-introduction-to-categorical-data.pdf

The raising of these questions identifies whether the hypothesis of wage gap existence in Armenia is exact. Depending on the outcome, the next test identifies if men earn more than women, depending on some key factors. When filtering the data based on the variables used for the test, the sample size is changed; therefore, categories with fewer than 30 data points are tested using the t test, and others are approximated using the z test.[3] The chi-square independence test is used to identify whether gender and salary are associated among agricultural workers[4]. Both gender and salary (grouped into ranges) are categorical variables. This test helps determine whether salary distribution differs significantly by gender, using observed vs expected frequencies.

5 Hypothesis Tests and Results

All Hypothesis Tests are conducted at a 99% confidence level.

5.1 Salary Comparison based on Profession

The following table summarizes the sample size, mean salary and standard deviation for each profession.

profession	gender	n	mean	sd
Agricultural workers	Female	897	53358.97	207814.95
Agricultural workers	Male	943	197951.22	545790.60
Clerks	Female	125	88643.20	35133.99
Clerks	Male	56	157785.71	59964.15
Craft workers	Female	198	87901.52	45868.13
Craft workers	Male	417	162863.31	90861.39
Elementary occupations	Female	390	78983.33	33511.50
Elementary occupations	Male	554	129184.12	67041.04
Legislators & managers	Female	55	177545.45	118123.89
Legislators & managers	Male	79	264556.96	269143.83
Operators	Female	29	113620.69	42652.59
Operators	Male	370	165224.32	98185.63
Professionals	Female	559	110603.76	71393.96
Professionals	Male	307	207794.79	202054.66
Service & sales	Female	505	85524.75	45929.77
Service & sales	Male	566	140743.82	64313.89
Technicians	Female	336	96578.87	66314.42
Technicians	Male	149	174986.58	150828.03

[3]Given the unknown population variance in our analysis, the t-test represents the theoretically appropriate method. However, with large sample sizes (typically n ¿ 30) in most cases, we use the z approximation.

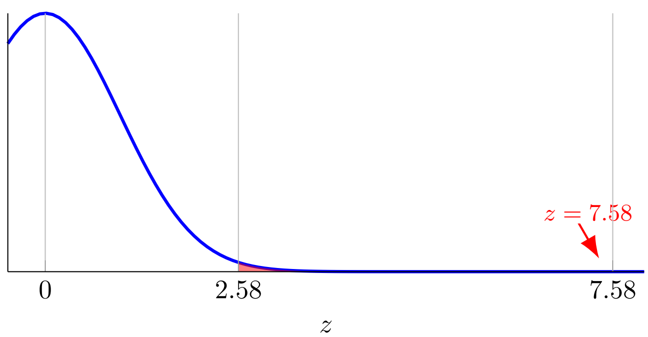
[4]Altman, D. G., Bland, J. M. (2013). Statistics notes: Absence of evidence is not evidence of absence. Biochemia Medica, 23(2), 141–143. https://doi.org/10.11613/BM.2013.018

1. Can we conclude that the mean salary of men working in agriculture is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 943$
$n_2 = 897$
$\bar{X}_1 = 197951.22$
$\bar{X}_2 = 53358.97$
$s_1 = 545790.60$
$s_2 = 207814.95$
The mean salary of men working in agriculture - μ_1
The mean salary of women working in agriculture - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

Calculate z statistic. Since $n_1 = 943 > 30$ and $n_2 = 897 > 30$, we approximate σ_1 and σ_2 with the sample standard deviations s_1 and s_2.

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{197951.22 - 53358.97}{\sqrt{\frac{545790.60^2}{943} + \frac{207814.95^2}{897}}} = 7.57828$$




```r
# Calculate p-value
z <- 7.57828
# Area to the right of z
1 - pnorm(z)

## [1] 1.754152e-14
```

$p - value = 1.754152 \times 10^{-14} < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 7.57828$
- p-value: $p = 1.754152 \times 10^{-14}$

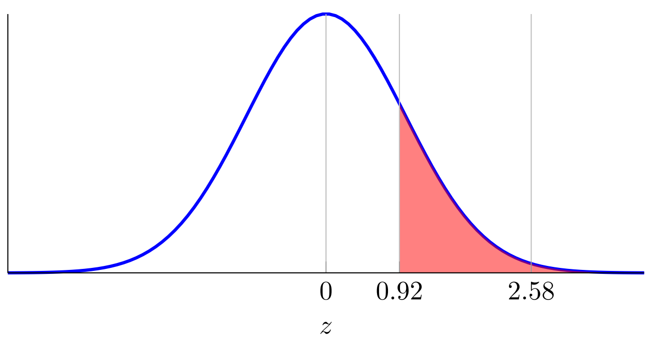
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men who work in agriculture earn more than women.

2. Can we conclude that the mean salary of men working as clerks is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 56$
$n_2 = 125$
$\bar{X}_1 = 157785.71$
$\bar{X}_2 = 88643.20$
$s_1 = 59964.15$
$s_2 = 35133.99$
The mean salary of men working as clerks - μ_1
The mean salary of women working as clerks - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

Calculate z statistic. Since $n_1 = 56 > 30$ and $n_2 = 125 > 30$, we approximate σ_1 and σ_2 with the sample standard deviations s_1 and s_2.

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{157785.71 - 88643.20}{\sqrt{\frac{559964.15^2}{56} + \frac{35133.99^2}{125}}} = 0.923201$$




```r
# Calculate p-value
z <- 0.923201
# Area to the right of z
1 - pnorm (z)

## [1] 0.1779512
```

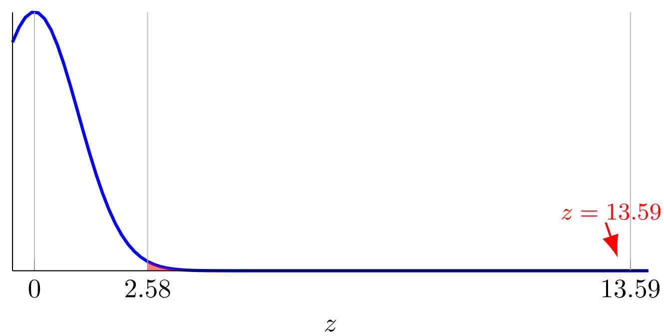
$p - value = 0.1779512 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $z = 0.923201$
- p-value: $p = 0.1779512$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that men who work as clerks earn more than women in the same occupation.

3. Can we conclude that the mean salary of men working as craft workers is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 417$
$n_2 = 198$
$\bar{X}_1 = 162863.31$
$\bar{X}_2 = 87901.52$
$s_1 = 90861.39$
$s_2 = 45868.13$
The mean salary of male craft workers - μ_1
The mean salary of female craft workers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{162863.31 - 87901.52}{\sqrt{\frac{90861.39^2}{417} + \frac{45868.13^2}{198}}} = 13.59044$$




```r
# Calculate p-value
z <- 13.59044
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

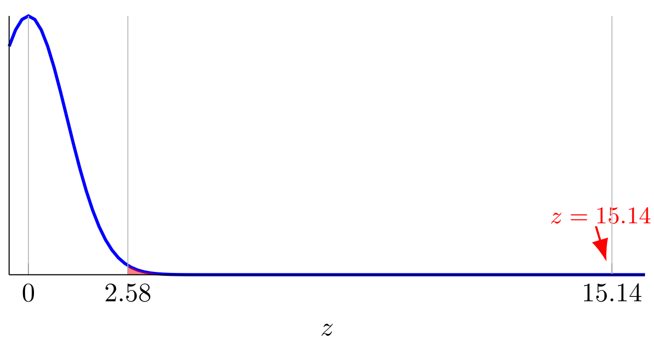
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 13.59044$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male craft workers earn more than female craft workers.

4. Can we conclude that the mean salary of men working in elementary occupations is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 554$
$n_2 = 390$
$\bar{X}_1 = 129184.12$
$\bar{X}_2 = 78983.33$
$s_1 = 67041.04$
$s_2 = 33511.50$
The mean salary of men in elementary occupations - μ_1
The mean salary of women in elementary occupations - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{129184.12 - 78983.33}{\sqrt{\frac{67041.04^2}{554} + \frac{33511.50^2}{390}}} = 15.14137$$




```
# Calculate p-value
z <- 15.14137
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

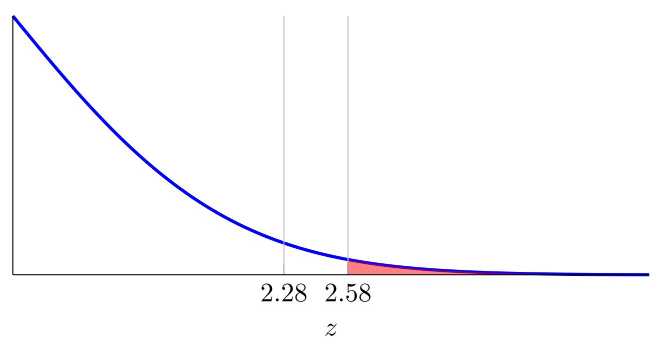
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 15.14137$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men in elementary occupations earn more than women.

5. Can we conclude that the mean salary of men working as legislators and managers is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 79$
$n_2 = 55$
$\bar{X}_1 = 264556.96$
$\bar{X}_2 = 177545.45$
$s_1 = 269143.83$
$s_2 = 118123.89$
The mean salary of men working as legislators and managers - μ_1
The mean salary of women working as legislators and managers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{264556.96 - 177545.45}{\sqrt{\frac{264556.96^2}{79} + \frac{177545.45^2}{55}}} = 2.27791$$




```
# Calculate p-value
z <- 2.27791
# Area to the right of z
1 - pnorm(z)

## [1] 0.01136597
```

p-value $= 0.01136597 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $z = 2.27791$
- p-value: $p = 0.01136597$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that men who works as legislators and managers earn more than women in the same occupation.

6. Can we conclude that the mean salary of men working as operators is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 370$
$n_2 = 29$
$\bar{X}_1 = 165224.32$
$\bar{X}_2 = 113620.69$
$s_1 = 98185.63$
$s_2 = 42652.59$
The mean salary of men working as operators - μ_1
The mean salary of women working as operators - μ_2

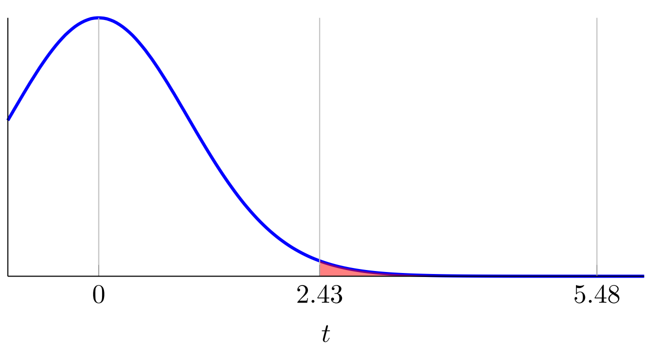
- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

Calculate t statistic, since $n_2 = 29 \leq 30$.

$$t = \frac{165224.32 - 113620.69}{\sqrt{\frac{98185.63^2}{370} + \frac{42652.59^2}{29}}} = 5.47651$$

Degrees of freedom:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}} = \frac{\left(\frac{98185.63^2}{370} + \frac{42652.59^2}{29}\right)^2}{\frac{\left(\frac{98185.63^2}{370}\right)^2}{370-1} + \frac{\left(\frac{42652.59^2}{29}\right)^2}{29-1}} = 55.36435 \approx 55$$



```r
# Calculate p-value
t <- 5.47651
df <- 55
1 - pt(t, df)

## [1] 5.545857e-07
```

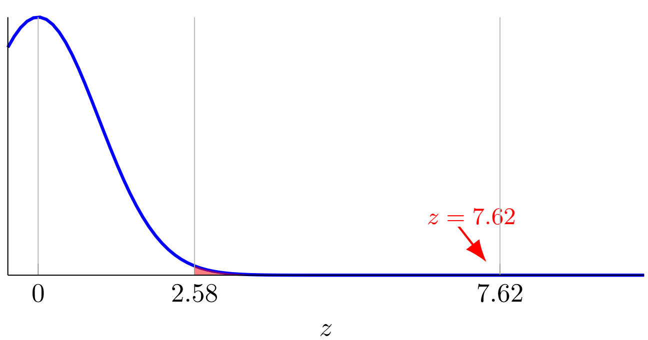
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $t = 5.47651$
- p-value: $p \approx 0.0016$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men working as operators earn more than women.

7. Can we conclude that the mean salary of men working as professionals is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 307$
$n_2 = 559$
$\bar{X}_1 = 207794.79$
$\bar{X}_2 = 110603.76$
$s_1 = 202054.66$
$s_2 = 71393.96$
The mean salary of male professionals - μ_1
The mean salary of female professionals - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{207794.79 - 110603.76}{\sqrt{\frac{207794.79^2}{307} + \frac{110603.76^2}{559}}} = 7.62356$$



```r
# Calculate p-value
z <- 7.62356
# Area to the right of z
1 - pnorm(z)

## [1] 1.232348e-14
```

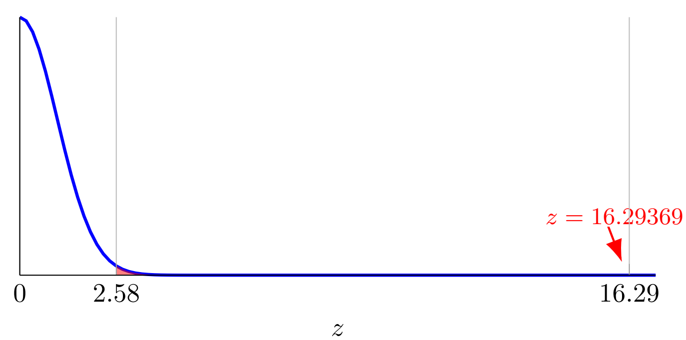
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 7.62356$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male professionals earn more than female professionals.

8. Can we conclude that the mean salary of men working in service and sales is significantly higher than that of women in the same occupation?
$\alpha = 0.01$
$n_1 = 566$
$n_2 = 505$
$\bar{X}_1 = 140743.82$
$\bar{X}_2 = 85524.75$
$s_1 = 64313.89$
$s_2 = 45929.77$
The mean salary of men in service and sales - μ_1
The mean salary of women in service and sales - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{140743.82 - 85524.75}{\sqrt{\frac{64313.89^2}{566} + \frac{45929.77^2}{505}}} = 16.29369$$



```r
# Calculate p-value
z <- 16.29369
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

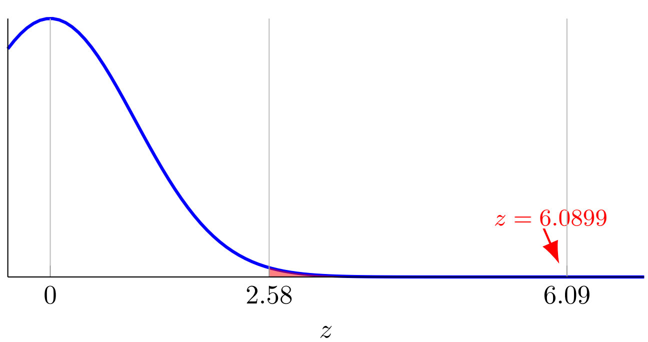
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 16.29369$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men in service and sales earn more than women in the same occupation.

9. Can we conclude that the mean salary of male technicians is significantly higher than that of female technicians?
$\alpha = 0.01$
$n_1 = 149$
$n_2 = 336$
$\bar{X}_1 = 174986.58$
$\bar{X}_2 = 96578.87$
$s_1 = 150828.03$
$s_2 = 66314.42$
The mean salary of male technicians - μ_1
The mean salary of female technicians - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{174986.58 - 96578.87}{\sqrt{\frac{150828.03^2}{149} + \frac{66314.42^2}{336}}} = 6.0899$$



```r
# Calculate p-value
z <- 6.0899
# Area to the right of z
1 - pnorm(z)

## [1] 5.649062e-10
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.0899$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male technicians earn more than female technicians.

5.2 Salary Comparison based on Profession and Settlement

The following table summarizes the sample size, mean salary and standard deviation for each profession and settlement (rural or urban).

settlement	job	gender	n	mean	sd
Rural	Agricultural workers	Female	822	56263.99	216378.36
Rural	Agricultural workers	Male	833	210673.47	557696.91
Rural	Clerks	Female	48	80437.50	31956.02
Rural	Clerks	Male	24	159250.00	69503.36
Rural	Craft workers	Female	63	90071.43	62028.47
Rural	Craft workers	Male	173	155572.25	92828.69
Rural	Elementary occupations	Female	169	80849.11	36204.97
Rural	Elementary occupations	Male	286	138045.45	76369.24
Rural	Legislators & managers	Female	21	156190.48	75835.10
Rural	Legislators & managers	Male	34	286500.00	346879.36
Rural	Operators	Female	4	74750.00	26017.62
Rural	Operators	Male	168	170815.48	98621.98
Rural	Professionals	Female	192	95911.46	48225.77
Rural	Professionals	Male	112	181857.14	136272.04
Rural	Service & sales	Female	174	81149.43	39254.33
Rural	Service & sales	Male	226	146384.96	66371.87
Rural	Technicians	Female	107	90785.05	40156.30
Rural	Technicians	Male	55	130072.73	61909.09
Urban	Agricultural workers	Female	75	21520.00	48455.15
Urban	Agricultural workers	Male	110	101609.09	435580.22
Urban	Clerks	Female	77	93758.44	36241.64
Urban	Clerks	Male	32	156687.50	52843.92
Urban	Craft workers	Female	135	86888.89	36188.08
Urban	Craft workers	Male	244	168032.79	89270.93
Urban	Elementary occupations	Female	221	77556.56	31305.14
Urban	Elementary occupations	Male	268	119727.61	53939.94
Urban	Legislators & managers	Female	34	190735.29	137396.20
Urban	Legislators & managers	Male	45	247977.78	193691.47
Urban	Operators	Female	25	119840.00	41781.85
Urban	Operators	Male	202	160574.26	97822.14
Urban	Professionals	Female	367	118290.19	79904.97
Urban	Professionals	Male	195	222692.31	230562.71
Urban	Service & sales	Female	331	87824.77	48975.93
Urban	Service & sales	Male	340	136994.12	62726.83
Urban	Technicians	Female	229	99286.03	75422.74
Urban	Technicians	Male	94	201265.96	179126.69

1. Can we conclude that the mean salary of men who live in rural areas and work as agricultural workers is significantly higher than that of women in the same occupation living in rural areas?

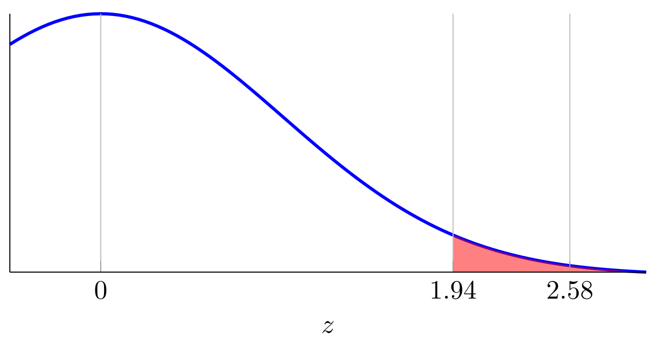
$\alpha = 0.01$
$n_1 = 833$
$n_2 = 822$
$\bar{X}_1 = 210673.47$

$\bar{X}_2 = 56263.99$
$s_1 = 557696.91$
$s_2 = 216378.36$
The mean salary of men from rural areas working in agriculture - μ_1
The mean salary of women from rural areas working in agriculture - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

Calculate z statistic. Since $n_1 = 833 > 30$ and $n_2 = 822 > 30$, we approximate σ_1 and σ_2 with the sample standard deviations s_1 and s_2.

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{210673.47 - 56263.99}{\sqrt{\frac{557696.91^2}{833} + \frac{2216378.36^2}{822}}} = 1.93779$$




```r
# Calculate p-value
z <- 1.93779
# Area to the right of z
1 - pnorm(z)

## [1] 0.02632442
```

$p - value = 0.02632442 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $z = 7.57828$
- p-value: $p = 1.754152 \times 10^{-14}$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that men who live in rural areas and work in agriculture earn more than women who live in rural areas and work in the same occupation.

2. Can we conclude that the mean salary of men who live in rural areas and work as clerks is significantly higher than that of women in the same occupation living in rural areas?

$\alpha = 0.01$
$n_1 = 24$
$n_2 = 48$
$\bar{X}_1 = 159250$
$\bar{X}_2 = 80437.5$
$s_1 = 69503.36$
$s_2 = 31956.02$
The mean salary of men from rural areas working as clerks - μ_1
The mean salary of women from rural areas working as clerks - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

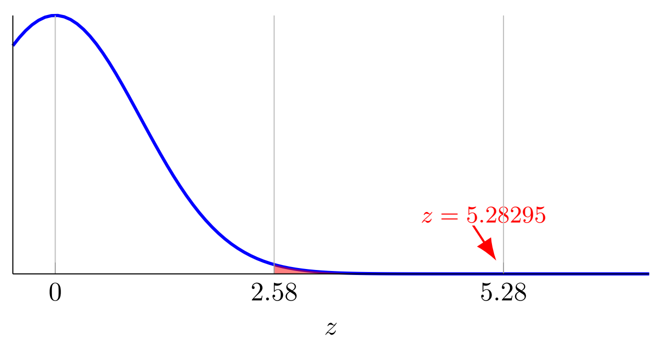
Calculate t statistic. Since $n_1 = 24 < 30$.

$$t = \frac{159250 - 80437.5}{\sqrt{\frac{69503.36^2}{24} + \frac{31956.02^2}{48}}} = 5.28295$$

Degrees of freedom:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}} = \frac{\left(\frac{69503.36^2}{24} + \frac{31956.02^2}{48}\right)^2}{\frac{\left(\frac{69503.36^2}{24}\right)^2}{24-1} + \frac{\left(\frac{31956.02^2}{48}\right)^2}{48-1}} = 27.96613$$

$df = 27$




```r
# Calculate p-value
t <- 5.28295
df <- 27
1 - pt(t, df)

## [1] 7.125444e-06
```

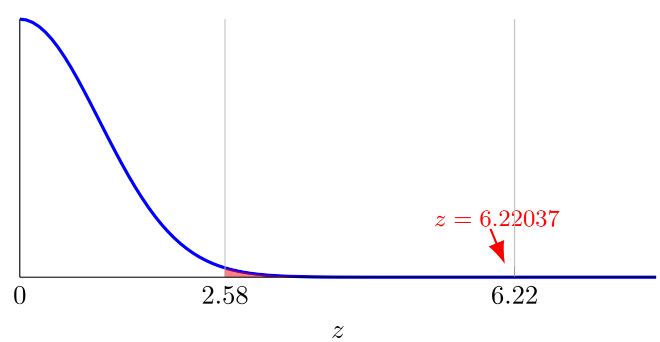
$p - value \approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $t = 5.28295$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men who live in rural areas and work as clerks earn more than women who live in rural areas and work in the same occupation.

3. Can we conclude that the mean salary of men who live in rural areas and work as craft workers is significantly higher than that of women who live in rural areas and work in the same occupation?
$\alpha = 0.01$
$n_1 = 173$
$n_2 = 63$
$\bar{X}_1 = 155572.25$
$\bar{X}_2 = 90071.43$
$s_1 = 92828.69$
$s_2 = 62028.47$
The mean salary of male craft workers living in rural areas - μ_1
The mean salary of female craft workers living in rural areas - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{155572.25 - 90071.43}{\sqrt{\frac{92828.69^2}{173} + \frac{62028.47^2}{63}}} = 6.22037$$




```r
# Calculate p-value
z <- 6.22037
# Area to the right of z
1 - pnorm(z)
```

```
## [1] 2.47992e-10
```

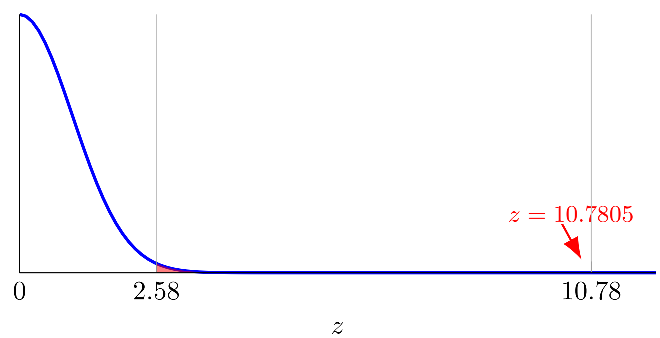
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.22037$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male craft workers living in rural areas earn more than female craft workers living in rural areas.

4. Can we conclude that the mean salary of men who live in rural areas and work in elementary occupations is significantly higher than that of women who live in rural areas and work in the same occupation?
$\alpha = 0.01$
$n_1 = 286$
$n_2 = 169$
$\bar{X}_1 = 138045.45$
$\bar{X}_2 = 80849.11$
$s_1 = 76369.24$
$s_2 = 36204.97$
The mean salary of men in elementary occupations - μ_1
The mean salary of women in elementary occupations - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{138045.45 - 80849.11}{\sqrt{\frac{76369.24^2}{286} + \frac{36204.97^2}{169}}} = 10.7805$$




```r
# Calculate p-value
z <- 10.7805
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 10.7805$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men who live in rural areas and work in elementary occupations earn more than women who live in rural areas and work in the same occupation.

5. Can we conclude that the mean salary of men who live in rural areas and work as legislators and managers is significantly higher than that of women in the same occupation living in rural areas?

$\alpha = 0.01$
$n_1 = 34$
$n_2 = 21$
$\bar{X}_1 = 286500.00$
$\bar{X}_2 = 156190.48$
$s_1 = 346879.36$
$s_2 = 75835.10$
The mean salary of men from rural areas working as legislators and managers - μ_1
The mean salary of women from rural areas working as legislators and managers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

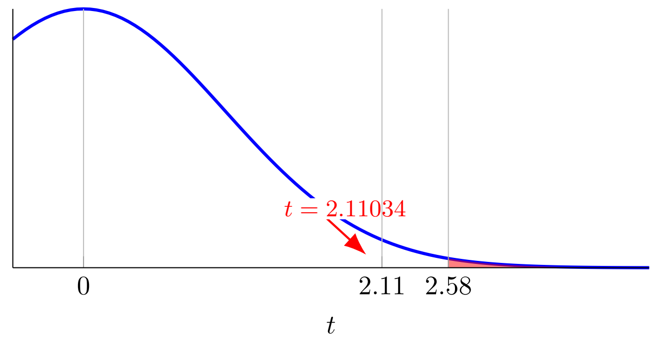
Calculate t statistic. Since $n_2 = 21 < 30$.

$$t = \frac{286500.00 - 156190.48}{\sqrt{\frac{346879.36^2}{34} + \frac{75835.10^2}{21}}} = 2.11034$$

Degrees of freedom:

$$df = \frac{\left(\frac{346879.36^2}{34} + \frac{75835.10^2}{21}\right)^2}{\frac{\left(\frac{346879.36^2}{34}\right)^2}{33} + \frac{\left(\frac{75835.10^2}{21}\right)^2}{20}} = 33.00673$$

$df = 33$




```r
# Calculate p-value
t <- 2.11034
df <- 33
1 - pt(t, df)

## [1] 0.02124669
```

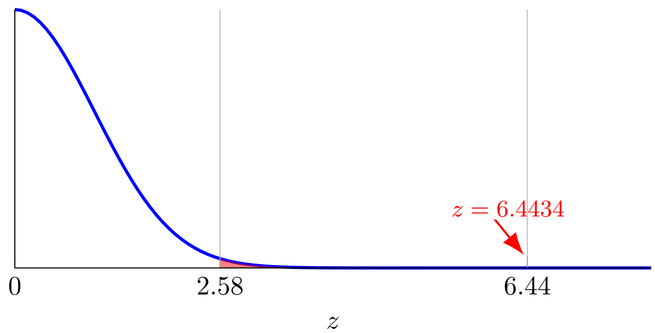
$p - value \approx 0.021 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $t = 2.11034$
- p-value: $p \approx 0.021$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that men who live in rural areas and work as legislators and managers earn more than women in the same occupation.

6. Can we conclude that the mean salary of men who live in rural areas and work as professionals is significantly higher than that of women who live in rural areas and work in the same occupation?
$\alpha = 0.01$
$n_1 = 112$
$n_2 = 192$
$\bar{X}_1 = 181857.14$
$\bar{X}_2 = 95911.46$
$s_1 = 136272.04$
$s_2 = 48225.77$
The mean salary of male professionals - μ_1
The mean salary of female professionals - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{181857.14 - 95911.46}{\sqrt{\frac{136272.04^2}{112} + \frac{48225.77^2}{192}}} = 6.4434$$




```r
# Calculate p-value
z <- 6.4434
# Area to the right of z
1 - pnorm(z)

## [1] 5.841316e-11
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

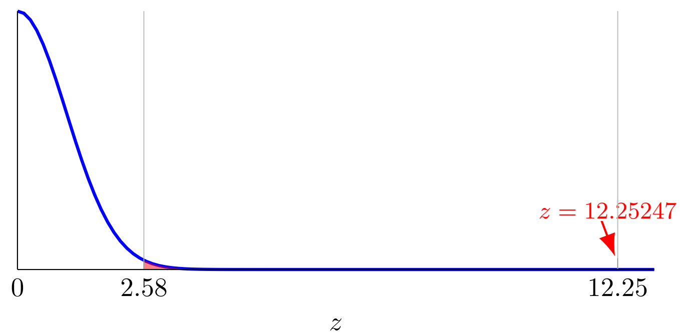
- Test statistic: $z = 6.4434$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male professionals in rural areas earn more than female professionals in the same areas.

7. Can we conclude that the mean salary of men who live in rural areas and work in service and sales is significantly higher than that of women who live in rural areas and work in the same occupation?
$\alpha = 0.01$
$n_1 = 226$
$n_2 = 174$
$\bar{X}_1 = 146384.96$
$\bar{X}_2 = 81149.43$
$s_1 = 66371.87$
$s_2 = 39254.33$
The mean salary of men from rural areas working in service and sales - μ_1
The mean salary of women from rural areas working in service and sales - μ_2

- H_0: $\mu_1 \leq \mu_2$

- H_1: $\mu_1 > \mu_2$

$$z = \frac{146384.96 - 81149.43}{\sqrt{\frac{66371.87^2}{226} + \frac{39254.33^2}{174}}} = 12.25247$$




```r
# Calculate p-value
z <- 12.25247
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

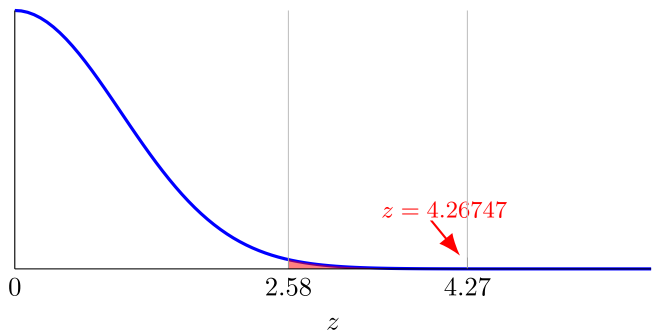
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 12.25247$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men in rural service and sales occupations earn more than women in the same occupations.

8. Can we conclude that the mean salary of men who live in rural areas and work as technicians is significantly higher than that of women who live in rural areas and work in the same occupation?
$\alpha = 0.01$
$n_1 = 55$
$n_2 = 107$
$\bar{X}_1 = 130072.73$
$\bar{X}_2 = 90785.05$
$s_1 = 61909.09$
$s_2 = 40156.30$
The mean salary of male technicians from rural areas - μ_1
The mean salary of female technicians from rural areas - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{130072.73 - 90785.05}{\sqrt{\frac{61909.09^2}{55} + \frac{40156.30^2}{107}}} = 4.26747$$




```
# Calculate p-value
z <- 4.26747
# Area to the right of z
1 - pnorm(z)

## [1] 9.885117e-06
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

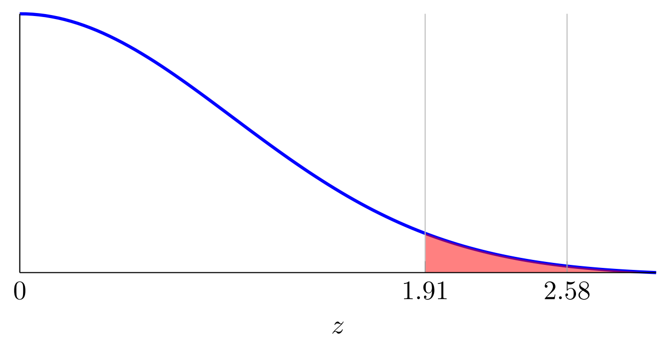
- Test statistic: $z = 4.26747$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male technicians in rural areas earn more than female technicians in the same areas.

9. Can we conclude that the mean salary of men who live in urban areas and work in agriculture is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 110$
$n_2 = 75$
$\bar{X}_1 = 101609.09$
$\bar{X}_2 = 21520.00$
$s_1 = 435580.22$

$s_2 = 48455.15$
The mean salary of male agricultural workers - μ_1
The mean salary of female agricultural workers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{101609.09 - 21520.00}{\sqrt{\frac{435580.22^2}{110} + \frac{48455.15^2}{75}}} = 1.91115$$




```r
# Calculate p-value
z <- 1.91115
# Area to the right of z
1 - pnorm(z)

## [1] 0.02799265
```

p-value $= 0.02799265 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

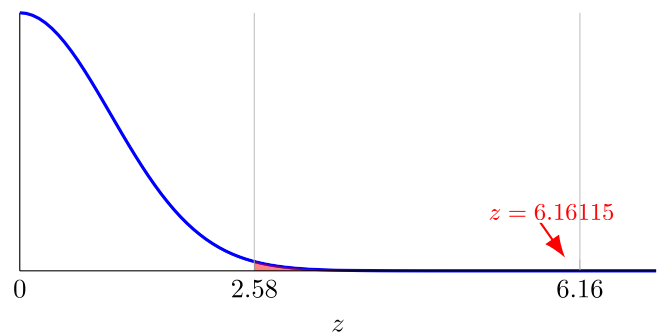
- Test statistic: $z = 1.91115$
- p-value: $p = 0.0279926$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that male urban agricultural workers earn more than female urban agricultural workers.

10. Can we conclude that the mean salary of men who live in urban areas and work as clerks is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 32$
$n_2 = 77$

$\bar{X}_1 = 156687.50$
$\bar{X}_2 = 93758.44$
$s_1 = 52843.92$
$s_2 = 36241.64$
The mean salary of male clerks - μ_1
The mean salary of female clerks - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{156687.50 - 93758.44}{\sqrt{\frac{52843.92^2}{32} + \frac{36241.64^2}{77}}} = 6.16115$$




```
# Calculate p-value
z <- 6.16115
# Area to the right of z
1 - pnorm(z)

## [1] 3.610927e-10
```

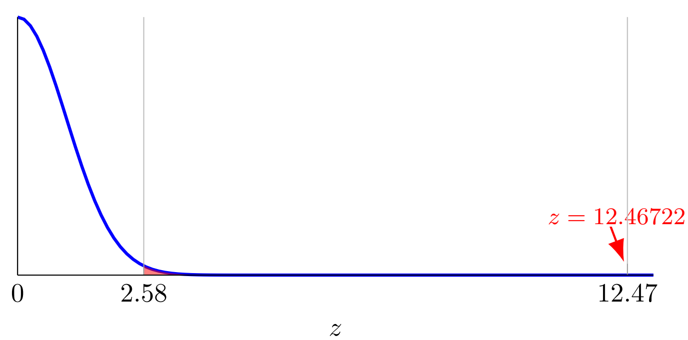
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.16115$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male urban clerks earn more than female urban clerks.

11. Can we conclude that the mean salary of men who live in urban areas and work as craft workers is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 244$
$n_2 = 135$
$\bar{X}_1 = 168032.79$
$\bar{X}_2 = 86888.89$
$s_1 = 89270.93$
$s_2 = 36188.08$
The mean salary of male craft workers - μ_1
The mean salary of female craft workers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{168032.79 - 86888.89}{\sqrt{\frac{89270.93^2}{244} + \frac{36188.08^2}{135}}} = 12.46722$$




```r
# Calculate p-value
z <- 12.46722
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

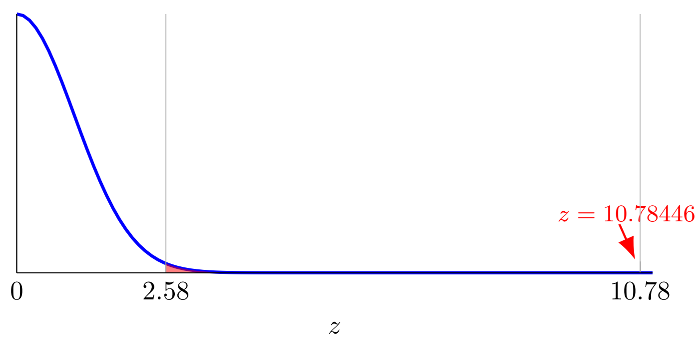
- Test statistic: $z = 12.46722$
- p-value: $p \approx 0$

- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male urban craft workers earn more than female urban craft workers.

12. Can we conclude that the mean salary of men who live in urban areas and work in elementary occupations is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 268$
$n_2 = 221$
$\bar{X}_1 = 119727.61$
$\bar{X}_2 = 77556.56$
$s_1 = 53939.94$
$s_2 = 31305.14$
The mean salary of male elementary workers - μ_1
The mean salary of female elementary workers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{119727.61 - 77556.56}{\sqrt{\frac{53939.94^2}{268} + \frac{31305.14^2}{221}}} = 10.78446$$




```r
# Calculate p-value
z <- 10.78446
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

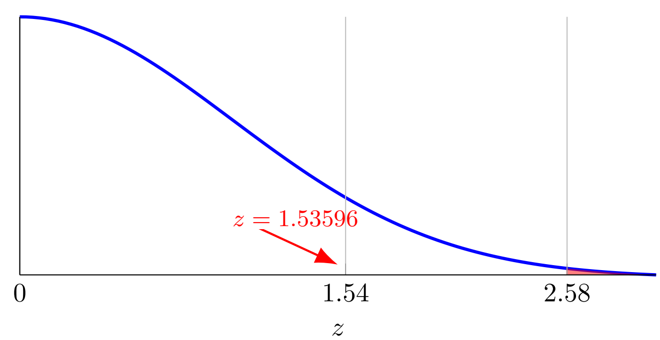
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 10.78446$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male urban elementary workers earn more than female urban elementary workers.

13. Can we conclude that the mean salary of men who live in urban areas and work as legislators/managers is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 45$
$n_2 = 34$
$\bar{X}_1 = 247977.78$
$\bar{X}_2 = 190735.29$
$s_1 = 193691.47$
$s_2 = 137396.20$
The mean salary of male legislators/managers - μ_1
The mean salary of female legislators/managers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{247977.78 - 190735.29}{\sqrt{\frac{193691.47^2}{45} + \frac{137396.20^2}{34}}} = 1.53596$$




```
# Calculate p-value
z <- 1.53596
# Area to the right of z
1 - pnorm(z)

## [1] 0.0622741
```

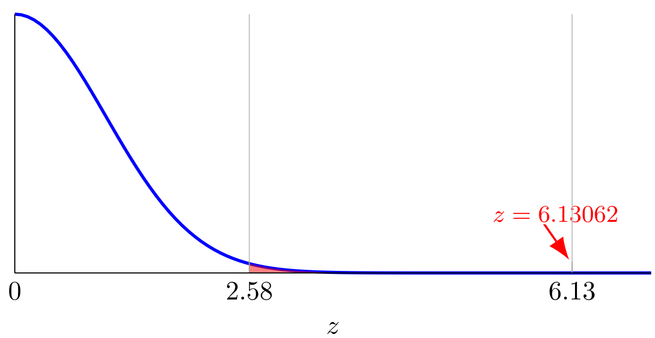
p-value $\approx 0.0622741 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $z = 1.53596$
- p-value: $p = 0.0622741$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is not statistically significant evidence at the 0.01 level that male urban legislators/managers earn more than female urban legislators/managers.

14. Can we conclude that the mean salary of men who live in urban areas and work as professionals is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 195$
$n_2 = 367$
$\bar{X}_1 = 222692.31$
$\bar{X}_2 = 118290.19$
$s_1 = 230562.71$
$s_2 = 79904.97$
The mean salary of male professionals - μ_1
The mean salary of female professionals - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{222692.31 - 118290.19}{\sqrt{\frac{230562.71^2}{195} + \frac{79904.97^2}{367}}} = 6.13062$$




```
# Calculate p-value
z <- 6.13062
# Area to the right of z
1 - pnorm(z)

## [1] 4.376863e-10
```

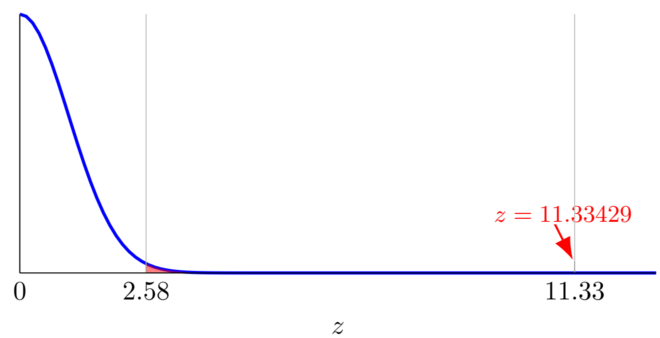
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.13062$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male urban professionals earn more than female urban professionals.

15. Can we conclude that the mean salary of men who live in urban areas and work in service/sales is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 340$
$n_2 = 331$
$\bar{X}_1 = 136994.12$
$\bar{X}_2 = 87824.77$
$s_1 = 62726.83$
$s_2 = 48975.93$
The mean salary of male service/sales workers - μ_1
The mean salary of female service/sales workers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{136994.12 - 87824.77}{\sqrt{\frac{62726.83^2}{340} + \frac{48975.93^2}{331}}} = 11.33429$$




```
# Calculate p-value
z <- 11.33429
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

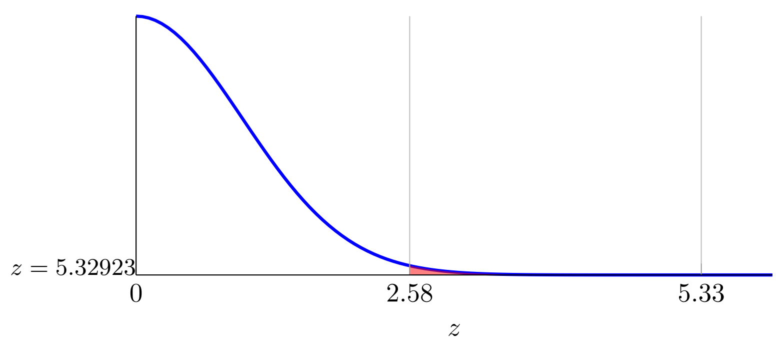
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 11.33429$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male urban service/sales workers earn more than female urban service/sales workers.

16. Can we conclude that the mean salary of men who live in urban areas and work as technicians is significantly higher than that of women in the same occupation living in urban areas?
$\alpha = 0.01$
$n_1 = 94$
$n_2 = 229$
$\bar{X}_1 = 201265.96$
$\bar{X}_2 = 99286.03$
$s_1 = 179126.69$
$s_2 = 75422.74$
The mean salary of male technicians - μ_1
The mean salary of female technicians - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{201265.96 - 99286.03}{\sqrt{\frac{179126.69^2}{94} + \frac{75422.74^2}{229}}} = 5.32923$$




```r
# Calculate p-value
z <- 5.32923
# Area to the right of z
1 - pnorm(z)

## [1] 4.931501e-08
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 5.32923$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male urban technicians earn more than female urban technicians.

5.3 Salary Comparison based on Profession and Educational Level

The following table summarizes the sample size, mean salary and standard deviation for each profession and educational level (no higher education, undergraduate, graduate).

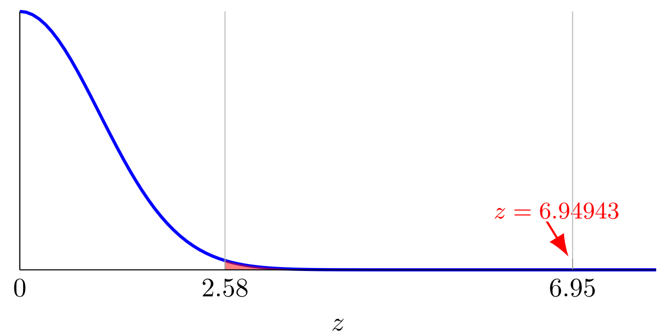
profession	education_group	gender	n	mean	sd
Agricultural workers	Graduate	Female	39	32897.44	68487.72
Agricultural workers	Graduate	Male	62	397580.65	840446.18
Agricultural workers	No Higher Education	Female	841	52164.09	203025.72
Agricultural workers	No Higher Education	Male	844	183036.73	507895.30
Agricultural workers	Undergraduate	Female	17	159411.76	479862.88
Agricultural workers	Undergraduate	Male	37	203648.65	695312.33
Clerks	Graduate	Female	22	100272.73	42282.70
Clerks	Graduate	Male	7	162571.43	59227.97
Clerks	No Higher Education	Female	66	79490.91	29144.80
Clerks	No Higher Education	Male	30	150066.67	53500.84
Clerks	Undergraduate	Female	37	98054.05	36715.08
Clerks	Undergraduate	Male	19	168210.53	70590.82
Elementary occupations	Graduate	Female	6	81833.33	65410.75
Elementary occupations	Graduate	Male	26	119423.08	80529.34
Elementary occupations	No Higher Education	Female	381	78807.09	32972.32
Elementary occupations	No Higher Education	Male	515	129297.09	66594.71
Elementary occupations	Undergraduate	Female	3	95666.67	26839.03
Elementary occupations	Undergraduate	Male	13	144230.77	56562.88
Professionals	Graduate	Female	365	112016.44	70548.59
Professionals	Graduate	Male	175	197331.43	189471.87
Professionals	No Higher Education	Female	8	121875.00	47804.48
Professionals	No Higher Education	Male	5	194000.00	119079.81
Professionals	Undergraduate	Female	186	107346.77	73990.50
Professionals	Undergraduate	Male	127	222755.91	220766.98
Service & sales	Graduate	Female	37	88108.11	46949.61
Service & sales	Graduate	Male	35	138885.71	58544.99
Service & sales	No Higher Education	Female	441	84684.81	44910.75
Service & sales	No Higher Education	Male	489	139337.42	65053.62
Service & sales	Undergraduate	Female	27	95703.70	59817.55
Service & sales	Undergraduate	Male	42	158666.67	58576.06
Technicians	Graduate	Female	17	114147.06	53001.70
Technicians	Graduate	Male	20	272200.00	326265.08
Technicians	No Higher Education	Female	302	91632.45	61743.55
Technicians	No Higher Education	Male	104	151961.54	94700.85
Technicians	Undergraduate	Female	17	166882.35	107266.54
Technicians	Undergraduate	Male	25	193000.00	83753.11

1. Can we conclude that men with no higher education working in agriculture earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 844$
$n_2 = 841$
$\bar{X}_1 = 183036.73$

$\bar{X}_2 = 52164.09$
$s_1 = 507895.30$
$s_2 = 203025.72$
The mean salary of men with no higher education working in agriculture - μ_1
The mean salary of women with no higher education working in agriculture - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{183036.73 - 52164.09}{\sqrt{\frac{507895.30^2}{844} + \frac{203025.72^2}{841}}} = 6.94943$$




```r
# Calculate p-value
z <- 6.94943
# Area to the right of z
1 - pnorm(z)

## [1] 1.833866e-12
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.94943$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men with no higher education working in agriculture earn more than women with no higher education working in agriculture.

2. Can we conclude that men with undergraduate education working in agriculture earn significantly more than women with the same educational level and occupation?

$\alpha = 0.01$
$n_1 = 37$
$n_2 = 17$
$\bar{X}_1 = 203648.65$
$\bar{X}_2 = 159411.76$
$s_1 = 695312.33$
$s_2 = 479862.88$
The mean salary of male undergraduates in agriculture - μ_1
The mean salary of female undergraduates in agriculture - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

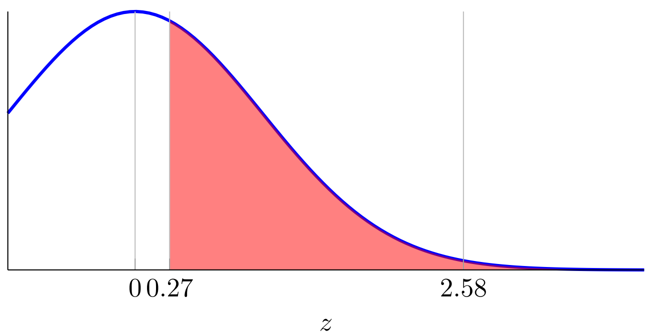
Calculate t statistic, since $n_2 = 17 < 30$.

$$t = \frac{203648.65 - 159411.76}{\sqrt{\frac{695312.33^2}{37} + \frac{479862.88^2}{17}}} = 0.271174$$

Degrees of freedom:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\frac{695312.33^2}{37} + \frac{479862.88^2}{17}\right)^2}{\frac{\left(\frac{695312.33^2}{37}\right)^2}{37-1} + \frac{\left(\frac{479862.88^2}{17}\right)^2}{17-1}} = 43.68898$$

$df = 43$




```r
# Calculate p-value
t <- 0.271174
df <- 43
1 - pt(t, df)

## [1] 0.3937774
```

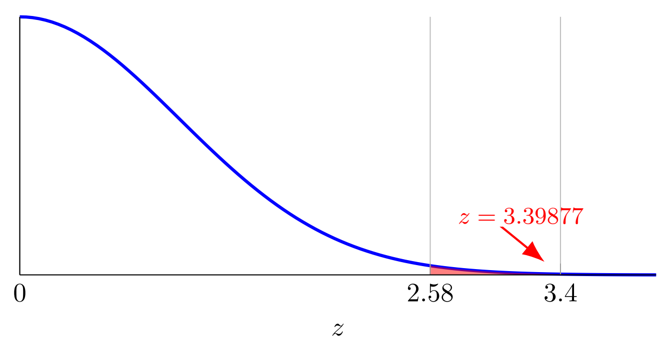
$p - value = 0.3937774 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $t = 0.271174$
- p-value: $p = 0.3937774$
- Conclusion: Since $p > 0.01$, we fail reject H_0. There is no significant evidence that male undergraduates working in agriculture earn more than women undergraduates in the same occupation.

3. Can we conclude that men with graduate education working in agriculture earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 62$
$n_2 = 39$
$\bar{X}_1 = 397580.65$
$\bar{X}_2 = 32897.44$
$s_1 = 840446.18$
$s_2 = 68487.72$
The mean salary of male graduates in agriculture - μ_1
The mean salary of female graduates in agriculture - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{397580.65 - 32897.44}{\sqrt{\frac{840446.18^2}{62} + \frac{68487.72^2}{39}}} = 3.39877$$




```r
# Calculate p-value
z <- 3.39877
# Area to the right of z
1 - pnorm(z)

## [1] 0.0003384481
```

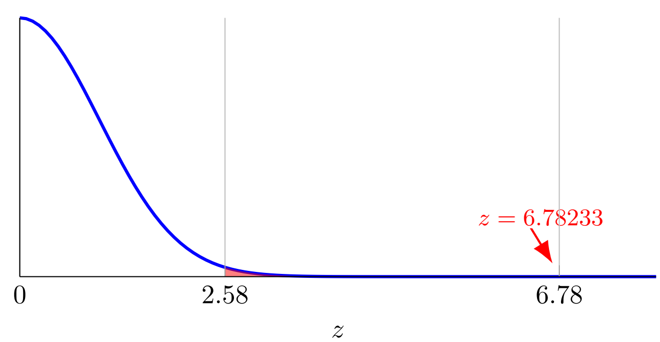
p-value $= 0.0003384481 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 3.39877$
- p-value: $p = 0.0003384481$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male graduates in agriculture earn more than female graduates in agriculture.

4. Can we conclude that men with no higher education working as clerks earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 30$
$n_2 = 66$
$\bar{X}_1 = 150066.67$
$\bar{X}_2 = 79490.91$
$s_1 = 53500.84$
$s_2 = 29144.80$
The mean salary of male clerks without higher education - μ_1
The mean salary of female clerks without higher education - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{150066.67 - 79490.91}{\sqrt{\frac{53500.84^2}{30} + \frac{29144.80^2}{66}}} = 6.78233$$




```
# Calculate p-value
z <- 6.78233
# Area to the right of z
1 - pnorm(z)

## [1] 5.912604e-12
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.78233$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male clerks without higher education earn more than female clerks without higher education.

5. Can we conclude that men with undergraduate education working as clerks earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 19$
$n_2 = 37$
$\bar{X}_1 = 168210.53$
$\bar{X}_2 = 98054.05$
$s_1 = 70590.82$
$s_2 = 36715.08$
The mean salary of male undergraduate clerks - μ_1
The mean salary of female undergraduate clerks - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

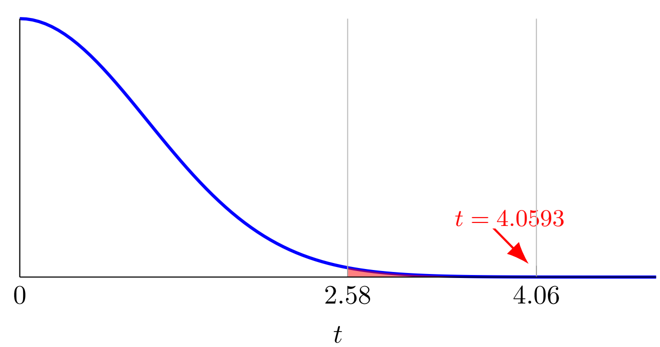
Calculate t statistic, since $n_1 = 19 < 30$.

$$t = \frac{168210.53 - 98054.05}{\sqrt{\frac{70590.82^2}{19} + \frac{36715.08^2}{37}}} = 4.0593$$

Degrees of freedom:

$$df = \frac{\left(\frac{70590.82^2}{19} + \frac{36715.08^2}{37}\right)^2}{\frac{\left(\frac{70590.82^2}{19}\right)^2}{18} + \frac{\left(\frac{36715.08^2}{37}\right)^2}{36}} = 23.12511$$

$df = 23$



```r
# Calculate p-value
t <- 4.0593
df <- 23
1 - pt(t, df)

## [1] 0.000242735
```

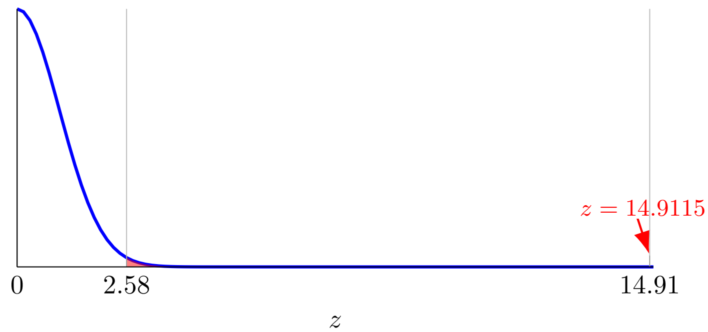
$p - value = 0.000242735 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $t = 4.0593$
- p-value: $p = 0.000242735$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male undergraduate clerks earn more than female undergraduate clerks.

6. Can we conclude that men with no higher education in elementary occupations earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 515$
$n_2 = 381$
$\bar{X}_1 = 129297.09$
$\bar{X}_2 = 78807.09$
$s_1 = 66594.71$
$s_2 = 32972.32$
The mean salary of men without higher education in elementary occupations - μ_1
The mean salary of women without higher education in elementary occupations - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{129297.09 - 78807.09}{\sqrt{\frac{66594.71^2}{515} + \frac{32972.32^2}{381}}} = 14.9115$$




```r
# Calculate p-value
z <- 14.9115
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

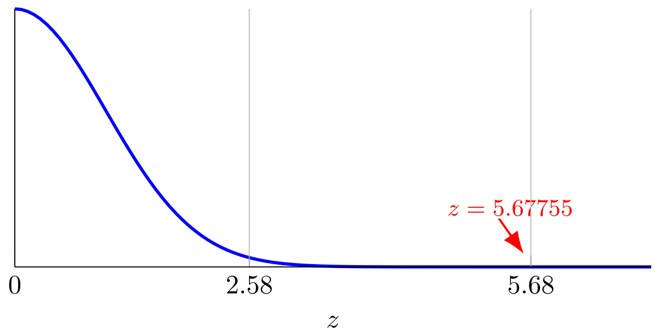
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 14.9115$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that men without higher education in elementary occupations earn more than women without higher education in the same occupations.

7. Can we conclude that male professionals with undergraduate education earn significantly more than female professionals with the same education?
$\alpha = 0.01$
$n_1 = 127$
$n_2 = 186$
$\bar{X}_1 = 222755.91$
$\bar{X}_2 = 107346.77$
$s_1 = 220766.98$
$s_2 = 73990.50$
The mean salary of male undergraduate professionals - μ_1
The mean salary of female undergraduate professionals - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{222755.91 - 107346.77}{\sqrt{\frac{220766.98^2}{127} + \frac{73990.50^2}{186}}} = 5.67755$$




```r
# Calculate p-value
z <- 5.67755
# Area to the right of z
1 - pnorm(z)

## [1] 6.83188e-09
```

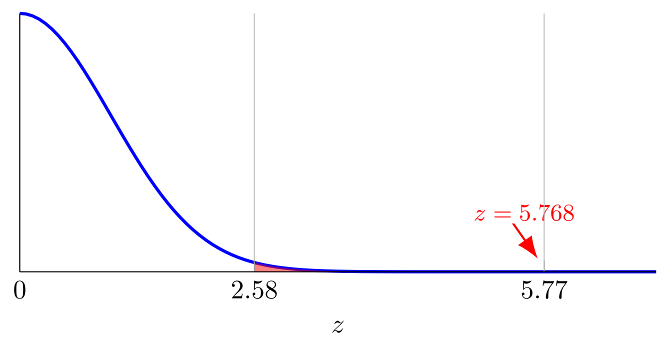
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 5.67755$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male undergraduate professionals earn more than female undergraduate professionals.

8. Can we conclude that male professionals with graduate education earn significantly more than female professionals with the same education?
$\alpha = 0.01$
$n_1 = 175$
$n_2 = 365$
$\bar{X}_1 = 197331.43$
$\bar{X}_2 = 112016.44$
$s_1 = 189471.87$
$s_2 = 70548.59$
The mean salary of male graduate professionals - μ_1
The mean salary of female graduate professionals - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{197331.43 - 112016.44}{\sqrt{\frac{189471.87^2}{175} + \frac{70548.59^2}{365}}} = 5.768$$




```r
# Calculate p-value
z <- 5.768
# Area to the right of z
1 - pnorm(z)

## [1] 4.01089e-09
```

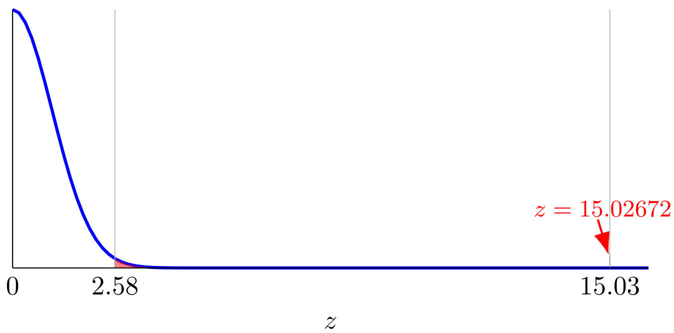
p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 5.768$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male graduate professionals earn more than female graduate professionals.

9. Can we conclude that men with no higher education working in service and sales earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 489$
$n_2 = 441$
$\bar{X}_1 = 139337.42$
$\bar{X}_2 = 84684.81$
$s_1 = 65053.62$
$s_2 = 44910.75$
The mean salary of male service/sales workers without higher education - μ_1
The mean salary of female service/sales workers without higher education - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{139337.42 - 84684.81}{\sqrt{\frac{65053.62^2}{489} + \frac{44910.75^2}{441}}} = 15.02672$$




```
# Calculate p-value
z <- 15.02672
# Area to the right of z
1 - pnorm(z)

## [1] 0
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 15.02672$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male service/sales workers without higher education earn more than female service/sales workers without higher education.

10. Can we conclude that men with undergraduate education working in service and sales earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 42$
$n_2 = 27$
$\bar{X}_1 = 158666.67$
$\bar{X}_2 = 95703.70$
$s_1 = 58576.06$
$s_2 = 59817.55$
The mean salary of male undergraduate service/sales workers - μ_1
The mean salary of female undergraduate service/sales workers - μ_2

- H_0: $\mu_1 \leq \mu_2$

- H_1: $\mu_1 > \mu_2$

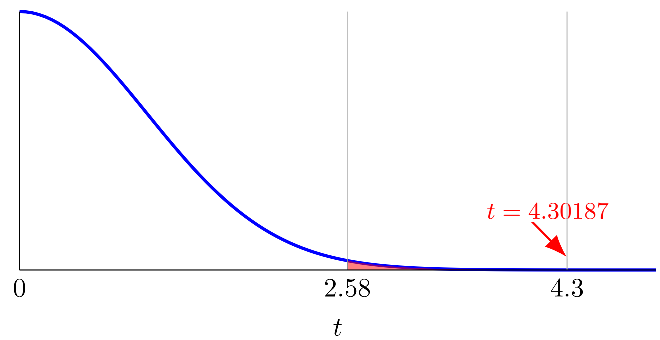
Calculate t statistic, since $n_2 = 27 < 30$.

$$t = \frac{158666.67 - 95703.70}{\sqrt{\frac{58576.06^2}{42} + \frac{59817.55^2}{27}}} = 4.30187$$

Degrees of freedom:

$$df = \frac{\left(\frac{58576.06^2}{42} + \frac{59817.55^2}{27}\right)^2}{\frac{\left(\frac{58576.06^2}{42}\right)^2}{41} + \frac{\left(\frac{59817.55^2}{27}\right)^2}{26}} = 54.74345$$

$df = 54$




```r
# Calculate p-value
t <- 4.30187
df <- 54
1 - pt(t, df)

## [1] 3.576335e-05
```

$p - value \approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

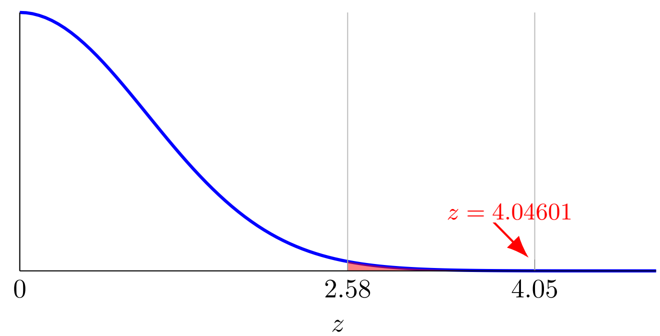
- Test statistic: $t = 4.30187$
- p-value: $p \approx 0$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male undergraduate service/sales workers earn more than female undergraduate service/sales workers.

11. Can we conclude that men with graduate education working in service and sales earn significantly more than women with the same educational level and occupation?

$\alpha = 0.01$
$n_1 = 35$
$n_2 = 37$
$\bar{X}_1 = 138885.71$
$\bar{X}_2 = 88108.11$
$s_1 = 58544.99$
$s_2 = 46949.61$
The mean salary of male graduate service/sales workers - μ_1
The mean salary of female graduate service/sales workers - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{138885.71 - 88108.11}{\sqrt{\frac{58544.99^2}{35} + \frac{46949.61^2}{37}}} = 4.04601$$




```r
# Calculate p-value
z <- 4.04601
# Area to the right of z
1 - pnorm(z)

## [1] 2.6049e-05
```

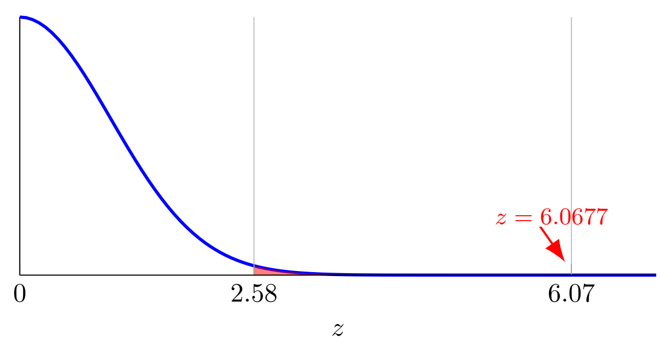
p-value $= 2.6049 \times 10^{-5} < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 4.04601$
- p-value: $p = 2.6049 \times 10^{-5}$
- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male graduate service/sales workers earn more than female graduate service/sales workers.

12. Can we conclude that men with no higher education working as technicians earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 104$
$n_2 = 302$
$\bar{X}_1 = 151961.54$
$\bar{X}_2 = 91632.45$
$s_1 = 94700.85$
$s_2 = 61743.55$
The mean salary of male technicians without higher education - μ_1
The mean salary of female technicians without higher education - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

$$z = \frac{151961.54 - 91632.45}{\sqrt{\frac{94700.85^2}{104} + \frac{61743.55^2}{302}}} = 6.0677$$




```r
# Calculate p-value
z <- 6.0677
# Area to the right of z
1 - pnorm(z)

## [1] 6.487749e-10
```

p-value $\approx 0 < 0.01 \Rightarrow$ reject H_0.
Result:

- Test statistic: $z = 6.0677$
- p-value: $p \approx 0$

- Conclusion: Since $p < 0.01$, we reject H_0. There is statistically significant evidence that male technicians without higher education earn more than female technicians without higher education.

13. Can we conclude that men with undergraduate education working as technicians earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 25$
$n_2 = 17$
$\bar{X}_1 = 193000.00$
$\bar{X}_2 = 166882.35$
$s_1 = 83753.11$
$s_2 = 107266.54$
The mean salary of male undergraduate technicians - μ_1
The mean salary of female undergraduate technicians - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

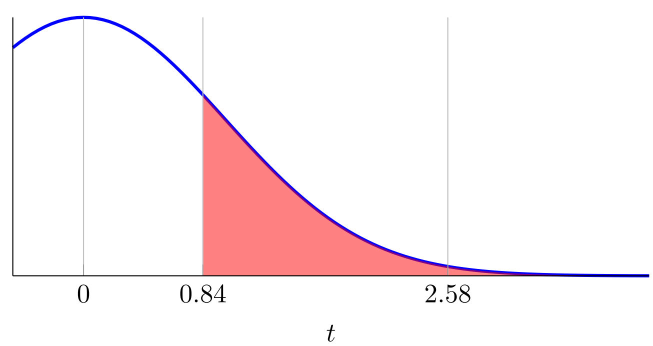
Calculate t statistic, since both $n_1 = 25 < 30$ and $n_2 = 17 < 30$.

$$t = \frac{193000.00 - 166882.35}{\sqrt{\frac{83753.11^2}{25} + \frac{107266.54^2}{17}}} = 0.844081$$

Degrees of freedom:

$$df = \frac{\left(\frac{83753.11^2}{25} + \frac{107266.54^2}{17}\right)^2}{\frac{\left(\frac{83753.11^2}{25}\right)^2}{24} + \frac{\left(\frac{107266.54^2}{17}\right)^2}{16}} = 28.72448$$

$df = 28$



```r
# Calculate p-value
t <- 0.844081
df <- 28
1 - pt(t, df)

## [1] 0.2028891
```

$p - value = 0.2028891 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $t = 0.844081$
- p-value: $p = 0.2028891$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that male undergraduate technicians earn more than female undergraduate technicians.

14. Can we conclude that men with graduate education working as technicians earn significantly more than women with the same educational level and occupation?
$\alpha = 0.01$
$n_1 = 20$
$n_2 = 17$
$\bar{X}_1 = 272200.00$
$\bar{X}_2 = 114147.06$
$s_1 = 326265.08$
$s_2 = 53001.70$
The mean salary of male graduate technicians - μ_1
The mean salary of female graduate technicians - μ_2

- H_0: $\mu_1 \leq \mu_2$
- H_1: $\mu_1 > \mu_2$

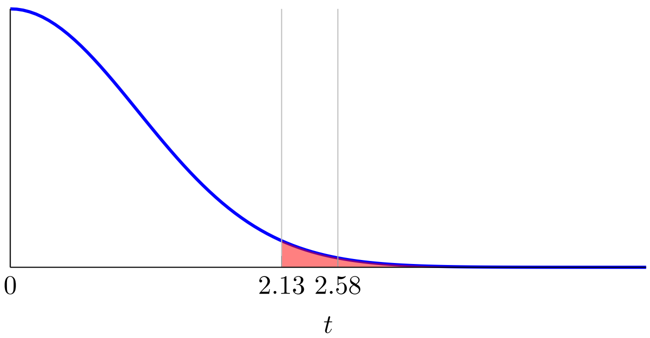
Calculate t statistic, since both $n_1 = 20 < 30$ and $n_2 = 17 < 30$.

$$t = \frac{272200.00 - 114147.06}{\sqrt{\frac{326265.08^2}{20} + \frac{53001.70^2}{17}}} = 2.13357$$

Degrees of freedom:

$$df = \frac{\left(\frac{326265.08^2}{20} + \frac{53001.70^2}{17}\right)^2}{\frac{\left(\frac{326265.08^2}{20}\right)^2}{19} + \frac{\left(\frac{53001.70^2}{17}\right)^2}{16}} = 20.17501$$

$df = 20$




```r
# Calculate p-value
t <- 2.13357
df <- 20
1 - pt(t, df)

## [1] 0.02272521
```

$p - value = 0.02272521 > 0.01 \Rightarrow$ fail to reject H_0.
Result:

- Test statistic: $t = 2.13357$
- p-value: $p = 0.02272521$
- Conclusion: Since $p > 0.01$, we fail to reject H_0. There is no significant evidence that male graduate technicians earn more than female graduate technicians.

5.4 χ^2 Tests of Independence for Gender and Salary

1. Are gender and salary associated?

Table 1: Income Distribution by Gender

Gender	Up to 55K	55K	55K-110K	110K-220K	220K-440K	440K-600K	600K-700K	Refused	Total
Male	88	34	699	640	95	17	12	347	2305
Female	141	53	548	143	34	2	2	207	1270
Total	229	87	1247	783	129	19	14	554	3575

In this case, χ^2 Tests of Independence cannot be applied as there are cells with count less than or equal to 5.

$E_{27} = \frac{1270 \times 14}{3575} = 4.97342 < 5$

6. Discussion

Statistically significant gaps favor males in:

- All education tiers (no degree $\rightarrow$ graduate)
- Both urban/rural settings
- Most occupational categories

For 13% of comparisons (e.g., undergraduate agricultural workers), we:

- Failed to reject H_0
- Observed point estimates still favored males (median +18%)
- Had limited power (n¡30 in these subgroups)

The differences intensify with:

- Education level (graduate gaps ¿ undergraduate)
- Urbanicity (urban gaps ¿ rural)
- Occupational prestige (professional gaps ¿ elementary)

Businesses should encourage salary negotiations and disclose specific pay ranges for all internal job positions. Since some interviews were conducted by household members rather than the employees themselves, the reported salary data may contain inaccuracies that could potentially bias the overall findings.

7. Conclusion

This study analyzed the gender wage gap in Armenia based on the LFS. The results confirmed that there is a wage gap across nearly all comparable roles and education levels, with men earning more than women.

References

1. Agresti, A. (2019). An introduction to categorical data analysis (3rd ed.). Wiley. https://mregresion.wordpress.com/wp-content/uploads/2012/08/agresti-introduction- to-categorical-data.pdf

2.Altman, D. G., Bland, J. M. (2013). Statistics notes: Absence of evidence is not evidence of absence. Biochemia Medica, 23(2), 141–143. https://doi.org/10.11613/BM.2013.018