



Testing the Independence Between Gender and T-shirt Prices

Is Gender Associated With the Price?



Introduction

Does gender affect clothing prices?

There is a popular belief that states, on average, women pay more for their clothing than men.

I would like to test this theory by looking at the prices of men's and women's t-shirts at three different stores.

By performing a chi-square test of independence, I will determine once and for all if gender affects the pricing of clothing.

The Sample



The stores I chose my samples from were Walmart, Urban Outfitters, and Forever 21. I chose a sample of 100 shirts from each store. 50 of which were men's shirts and the other 50 were women's shirts.



T-shirts

T-shirts were defined as shirts with the following:

- Short sleeves
- Lack of a collar
- No buttons or graphics
- Covers the midriff
- Of cotton blend material
- Medium sized



H_0 : Gender has no bearing on the prices of clothing. Neither men nor women are paying more for t-shirts since gender does not influence the prices.

H_A : Gender and the prices of t-shirts are associated. Gender plays a role in the prices of men's and women's t-shirts.

I am looking to see whether gender is associated with the pricing of t-shirts. My null and alternative hypotheses are stated above.

STORE	PRICE RANGE					TOTAL
	\$0 - \$10	\$10+ - \$20	\$20+ - \$30	\$30+ - \$40	\$40+	
Walmart (male)	46	4	0	0	0	50
Urban Outfitters (male)	0	2	13	16	19	50
Forever 21 (male)	29	20	1	0	0	50
TOTAL	75	26	14	16	19	150
Walmart (female)	48	2	0	0	0	50
Urban Outfitters (female)	0	30	11	5	4	50
Forever 21 (female)	22	27	1	0	0	50
TOTAL	70	59	12	5	4	150

Collecting Data

Assumptions and Conditions

Are sample groups independent?

Yes.

It is reasonable to assume this test is independent since one shirt has no bearing on the price of another.

Also, $n < 10\%$ of the entire population of t-shirts.

Are the data counts?

Yes.

None of the data collected was quantitative. All data comprised of categorical counts.

Is the data sufficiently large?

No...

Not all expected counts in each cell were calculated to be ≥ 5 .

Expected cell counts are to be greater than or equal to 5

STORE	PRICE RANGE					TOTAL
	\$0 - \$10	\$10+ - \$20	\$20+ - \$30	\$30+ - \$40	\$40+	
Walmart (male)	46	4	0	0	0	50
Urban Outfitters (male)	25	9	5	5	6	50
Forever 21 (male)	29	20	1	0	0	50
TOTAL	75	26	14	16	19	150
Walmart (female)	48	2	0	0	0	50
Urban Outfitters (female)	23	20	4	2	1	50
Forever 21 (female)	22	27	1	0	0	50
TOTAL	70	59	12	5	4	150

Collecting Data (Expected Counts)



Running the Test

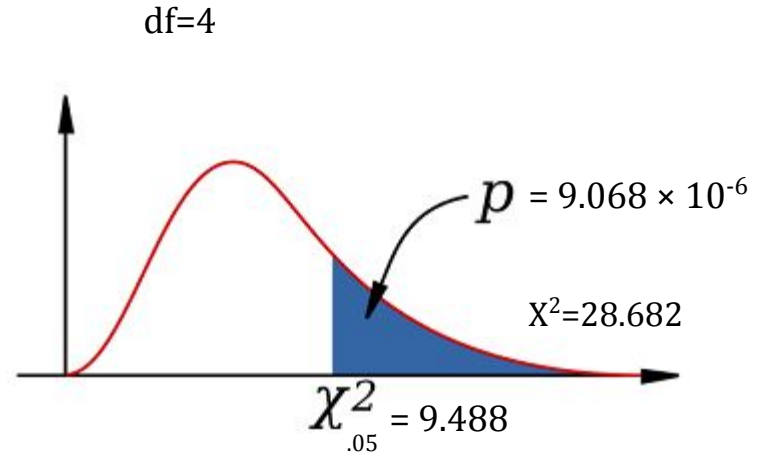
Using a 5% level of significance, is there enough evidence to support the alternative hypothesis?

I will continue on to run a chi-square test of independence, even though not all of my assumptions are met.

Once I determine the p-value and the chi-square value, I will be able to reject or fail to reject the null hypothesis.

Results

The results suggest that I should REJECT the null hypothesis.



Decision:
REJECT THE NULL



Conclusion

Since the calculated p-value (9.068×10^{-6}) was significantly lower than my 5% level of significance, I rejected the null hypothesis. This suggests that there is strong evidence to prove that gender and t-shirt prices are associated. Furthermore, the chi-square value (28.682) surpassed the chi-square boundary value of 9.488, so it is reasonable to reject the null hypothesis. The t-shirts from Urban Outfitters contributed the most the chi-square statistic, due to the fact that UO sells name brand clothing (ex. Ralph Lauren & FILA) that also fit the t-shirt requirements.

Men and women are not paying equally for clothing. It is possible that gender heavily affects the outcome of clothing prices, and one gender is spending more for clothing than the other. In other words, your gender may affect how much you pay for your clothing.



Limitations and Errors

It is possible for Type I and II errors to occur. A Type I error would involve me mistakenly stating that gender and t-shirt prices are associated when they are not.

A Type II error would involve me stating that gender and t-shirt prices are independent, when they are associated.

Neither has devastating consequences, but could alter the way we approach clothing stores.

My study was limited to three samples of 100 shirts from different stores and failed to meet all of the assumptions and conditions for a chi-square test. So, my conclusion may be a Type I error.

To improve the study, one might consider increasing the sample size to ensure the sample is sufficiently large.