Binomial goodness–of–fit test

In the next few slides I will work through the prize question at the end of Chapter 4 (Semester 2). Before you read through these slides, make sure you've got the question in front of you (question 3, page 39).

Appropriate distribution: **Binomial** distribution. Let X: Number of trains thrown away.

$$\begin{aligned} H_0 &: \text{ No. thrown away follows a Binomial distribution} \\ H_1 &: \text{ No. thrown away } \textit{doesn't} \text{ follow a Binomial!} \end{aligned}$$

We need to use the formula for the binomial distribution to calculate some probabilities (specifically, for $X = 0, 1, 2, 3, 4, 5$).

Binomial goodness–of–fit test

The binomial formula is:

$$\Pr(X = r) = {}^{n}\mathrm{C}_r \times p^r \times (1-p)^{n-r},$$

and we know that $n = 5$ and $p = 0.25$ (see the question). So we calculate

$$\begin{aligned}\Pr(X = 0) &= {}^{5}\mathrm{C}_0 \times 0.25^0 \times 0.75^5 \\ &= 0.2373\end{aligned}$$

We need to perform similar calculations for $X = 1, 2, 3, 4$ and 5. The results are shown in the table on the next slide. *[Note: We did not need to estimate ANY parameters here – both n and p were given to us!]*

Binomial goodness–of–fit test

We now have

X	Probability	Expected frequency
0	0.2373	12.3396
1	0.3955	20.5660
2	0.2634	13.6968
3	0.0879	4.5708
4	0.0146	0.7592
5	0.0010	0.0520

[Hint: multiply the probabilities by the sum of the frequencies – 52 here – to get the Expected frequencies!]

Notice we have three expected frequencies less than 5 – so we need to pool these categories, then we can calculate the test statistic!

X	Expected	Observed	$\frac{(O-E)^2}{E}$
0	12.3396	10	0.4436
1	20.5660	21	0.0092
2	13.6968	14	0.0067
3–5	5.3820	7	0.4864
			0.9459

Binomial goodness–of–fit test

The test statistic is $X^2 = 0.9459$. We need to compare this to tables. Now

$$\begin{aligned}\nu &= \text{No. of categories – no. of estimated parameters} - 1 \\ &= 4 - 0 - 1 = 3\end{aligned}$$

[Although the binomial distribution has two parameters, n and p, WE didn't need to estimate any of them – they were both given to us!]

From tables, using $\nu = 3$, we get:

p–value	10%	5%	1%
critical val.	6.25	7.82	11.34

So our *p*–value is bigger than 10%.

Binomial goodness–of–fit test

Conclusions

- No evidence against H_0
- Retain H_0
- The number of defective trains *does* follow a binomial distribution!