Chapter 24, Review Problem 7: Remember that the *error box* is associated with the measurement procedure, so the standard deviation of the error box is best estimated by the standard deviation of the (earlier) several hundred measurements. I.e., $SD \approx 18$ micrograms.
It follows that the standard error for the weight of the (new) 1 kg weight is

$$SE \approx \frac{18}{\sqrt{50}} \approx 2.545 \text{ micrograms},$$

and the confidence interval is therefore

$$(78.1 \pm 5.09) \text{ micrograms above 1 kg.}$$

Chapter 26, Review Problem 1:

(a) **True:** 'P-value' and 'observed significance level' are two ways of saying the same thing.

(b) **False:** The null hypothesis is the claim that the results are due to chance variation. The alternative hypothesis says that the results are ***not*** due to chance.

Chapter 26, Review Problem 8:
Box Model: there is one ticket in the box for each person in the county, age 18 and over. The ticket shows that persons educational level. The data are like 1000 draws from the box. The average of the box is μ.
Hypotheses:

- $H_0 : \mu = 13$
- $H_A : \mu \neq 13$

Standard error: $SE = \dfrac{\text{County SD}}{\sqrt{1000}} \approx \dfrac{\text{Sample SD}}{\sqrt{1000}} = \dfrac{5}{\sqrt{1000}} \approx 0.16$

Test statistic: $z = \dfrac{\text{(observed - expected)}}{SE} = \dfrac{14 - 13}{0.16} \approx 6.$

P-value: $p \approx$ area under normal curve outside of $(-6, 6) \approx 0$

Conclusion: The difference between the county sample average and the national average is not explained by chance sample variation. The average education level in the county is different (higher) than the national average. Perhaps this is a wealthy suburban county, or perhaps there is a college or university in the county — something that raises the average educational level in the county.

Chapter 27, Review Problem 3:

(a) This requires a two-sample z-test because we are comparing percentages from two different boxes: the 2000 box and the 2005 box.

(b) Box model: There are two 0-1 boxes, one for the year 2000 and one for the year 2005. In both boxes there is a $\boxed{1}$ for every person in the population (that year) who rates clergy as 'high' or 'very high' and a $\boxed{0}$ for every other person. The surveys are like simple random samples from each of these two boxes.

Hypotheses:

- H_0: $\%_{2000} = \%_{2005}$
- H_1 : $\%_{2000} > \%_{2005}$

This is a two-box test and $\%_{2000}$ and $\%_{2005}$ are the percentages of $\boxed{1}$s in 2000 and 2005, respectively. We use a one-sided test here because we are testing to see if sex scandals reduced the proportion of people who rate the clergy highly. The null hypothesis says that the percentage of people who rate clergy highly has not changed — the (observed) difference in sample percentages is due to chance. The alternative hypothesis says that the observed difference is due to a difference in the boxes.

(c) $SE_{2000} = \dfrac{\sqrt{0.6 \times 0.4}}{\sqrt{1000}} \times 100\% \approx 1.55\%$, $SE_{2005} = \dfrac{\sqrt{0.54 \times 0.46}}{\sqrt{1000}} \times 100\% \approx 1.58\%$ and

$$SE_{\text{diff}} = \sqrt{SE_{2000}^2 + SE_{2005}^2} \approx 2.21\%.$$

The z-score is

$$z = \frac{60\% - 54\%}{2.21\%} \approx 2.7,$$

and the P-value is approximately equal to the area under the normal curve to the right of $z = 2.7$, which is $p \approx 0.35\%$.

Conclusion: The difference is (almost certainly) ***not*** due to chance. On the other hand, we ***cannot*** tell from the percentage data what did cause the percentage to go down — whether it was sex scandals or something else.

Comment: A two-sided test could also be used here, and the conclusions would be the same. I.e., if the investigators did not have an *a priori* expectation that support for the clergy had gone down, they would (should) have used a two-sided alternative. This would have doubled the P-value from $p \approx 0.35\%$ to $p \approx 0.7\%$, which is still tiny.