Exam 2 Practice Questions –solutions, 18.05, Spring 2014

1 Topics

- Statistics: data, MLE (pset 5)
- Bayesian inference: prior, likelihood, posterior, predictive probability, probability intervals (psets 5, 6)
- Frequentist inference: NHST (psets 7, 8)

2 Using the probability tables

You should become familiar with the probability tables at the end of these notes.

1. (a) (i) The table gives this value as $P(Z < 1.5) = 0.9332$.

(ii) This is the complement of the answer in (i): $P(Z > 1.5) = 1 - 0.9332 = 0.0668$. Or by symmetry we could use the table for -1.5.

(iii) We want $P(Z < 1.5) - P(Z < -1.5) = P(Z < 1.5) - P(Z > 1.5)$. This is the difference of the answers in (i) and (ii): .8664.

(iv) A rough estimate is the average of $P(Z < 1.6)$ and $P(Z < 1.65)$. That is,

$$P(Z < 1.625) \approx \frac{P(Z < 1.6) + P(Z < 1.65)}{2} = \frac{.9452 + .9505}{2} = .9479.$$

(b) (i) We are looking for the table entry with probability 0.95. This is between the table entries for $z = 1.65$ and $z = 1.60$ and very close to that of $z = 1.65$. Answer: the region is $[1.64, \infty)$. (R gives the 'exact' lower limit as 1.644854.)

(ii) We want the table entry with probability 0.1. The table probabilities for $z = -1.25$ and $z = -1.30$ are 0.1056 and 0.0968. Since 0.1 is about 1/2 way from the first to the second we take the left critical value as -1.275. Our region is

$$(-\infty, -1.275) \cup (1.275, \infty).$$

(R gives `qnorm(0.1, 0, 1) = -1.2816`.)

(iii) This is the range from $q_{0.25}$ to $q_{0.75}$. With the table we estimate $q_{0.25}$ is about 1/2 of the way from -0.65 to -0.70, i.e. ≈ -0.675. So, the range is $[-0.675, 0.675]$.

2. (a) (i) The question asks for the table value for $t = 1.6$, $df = 3$: $\boxed{0.8960}$.

(ii) $P(T > 1.6) = 1 - P(T < 1.6) = 0.0703$.

(iii) By symmetry $P(T < -1.6) = 1 - P(T < 1.6)$. We want $P(T < 1.6) - P(T < -1.6) = 2P(T < 1.6) - 1 = 2(0.9420) - 1 = 0.8840$.

(iv) A rough estimate is the average of $P(T < 1.6)$ and $P(T < 1.8)$. That is,

$$P(T < 1.7) \approx \frac{P(T < 1.6) + P(T < 1.8)}{2} = \frac{0.9413 + 0.9603}{2} = 0.9508.$$

(The 'true' answer using R is 0.9516)

(b) (i) We are looking for the table entry with probability 0.95. This is about 1/3 of way between the table entries for $t = 1.8$ and $t = 2.0$ in the $df = 8$ column. Answer: 1.867. (R gives the 'exact' critical value as 1.860.)

(ii) We want the table entry with probability 0.9 in the $df = 16$ column. This is about 70% of the way from $t = 1.2$ to $t = 1.4$: $t = 1.34$. Our region is

$$(-\infty, -1.34) \cup (1.34, \infty).$$

(R gives `qt(0.1, 16) = -1.3368`.)

(iii) This is the range from $q_{0.25}$ to $q_{0.75}$. With the table we estimate $q_{0.75}$ is about 1/2 of the way from 0.6 to 0.8, i.e. ≈ -0.7. So, the range is $[-0.7, 0.7]$. (True answer is $[-0.6870, 0.6870]$.)

3. **(a)** (i) Looking in the $df = 3$ row of the chi-square table we see that 1.6 is about 2/2 of the way between the values for $p = 0.7$ and $p = .06$. So we approximate $P(X^2 > 1.6) \approx 0.66$. (The true value is 0.6594.)

(ii) Looking in the $df = 16$ row of the chi-square table we see that 20 is between the values for $p = 0.2$ and $p = 0.3$. We estimate $P(X^2 > 20) = 0.225$. (The true value is 0.220)

(b) (i) This is in the table in the $df = 8$ row under $p = 0.05$. Answer: 15.51

(ii) We want the critical values for $p = 0.9$ and $p = 0.1$ from the $df = 16$ row of the table.

$$[0, 9.31] \cup [23.54, \infty).$$

3 Data

4. Sample mean $20/5 = 4$.

Sample variance $= \dfrac{1^2 + (-3)^2 + (-1)^2 + (-1)^2 + 4^2}{5 - 1} = 7$.

Sample standard deviation $= \sqrt{7}$.

Sample median $= 3$.

4 MLE

5. (a) The likelihood function is

$$p(\text{data}|\theta) = \binom{100}{62}\theta^{62}(1-\theta)^{38} = c\theta^{62}(1-\theta)^{38}.$$

To find the MLE we find the derivative of the log-likelihood and set it to 0.

$$\ln(p(\text{data}|\theta)) = \ln(c) + 62\ln(\theta) + 38\ln(1-\theta).$$

$$\frac{d\,\ln(p(\text{data}|\theta))}{d\theta} = \frac{62}{\theta} - \frac{38}{1-\theta} = 0.$$

The algebra leads to the MLE $\boxed{\theta = 62/100}$.

(b) The computation is identical to part (a). The likelihood function is

$$p(\text{data}|\theta) = \binom{n}{k}\theta^{n}(1-\theta)^{k} = c\theta^{n}(1-\theta)^{k}.$$

To find the MLE we set the derivative of the log-likelihood and set it to 0.

$$\ln(p(\text{data}|\theta)) = \ln(c) + n\ln(\theta) + k\ln(1-\theta).$$

$$\frac{d\,\ln(p(\text{data}|\theta))}{d\theta} = \frac{n}{\theta} - \frac{k}{1-\theta} = 0.$$

The algebra leads to the MLE $\boxed{\theta = k/n}$.

6. If $N < \max(y_i)$ then the likelihood $p(y_1,\ldots,y_n|N) = 0$. So the likelihood function is

$$p(y_1,\ldots,y_n|N) = \begin{cases} 0 & \text{if } N < \max(y_i) \\ \left(\frac{1}{N}\right)^n & \text{if } N \geq \max(y_i) \end{cases}$$

This is maximized when N is as small as possible. Since $N \geq \max(y_i)$ the MLE is $\boxed{N = \max(y_i).}$

7. The pdf of $\exp(\lambda)$ is $p(x|\lambda) = \lambda \mathrm{e}^{-\lambda x}$. So the likelihood and log-likelihood functions are

$$p(\text{data}|\lambda) = \lambda^n \mathrm{e}^{-\lambda(x_1+\cdots+x_n)}, \qquad \ln(p(\text{data}|\lambda)) = n\ln(\lambda) - \lambda\sum x_i.$$

Taking a derivative with respect to λ and setting it equal to 0:

$$\frac{d\,\ln(p(\text{data}|\lambda))}{d\lambda} = \frac{n}{\lambda} - \sum x_i = 0 \qquad \Rightarrow\ \frac{1}{\lambda} = \frac{\sum x_i}{n} = \bar{x}.$$

So the MLE is $\boxed{\lambda = 1/\bar{x}}$.

8. $P(x_i|a) = \left(1 - \frac{1}{a}\right)^{x_i - 1} \frac{1}{a}. = \left(\frac{a-1}{a}\right)^{x_i-1} \frac{1}{a}.$

So, the likelihood function is

$$P(\text{data}|a) = \left(\frac{a-1}{a}\right)^{\sum x_i - n} \left(\frac{1}{a}\right)^n$$

The log likelihood is

$$\ln(P(\text{data}|a)) = \left(\sum x_i - n\right)(\ln(a-1) - \ln(a)) - n\ln(a).$$

Taking the derivative

$$\frac{d\,\ln(P(\text{data}|a))}{d\,a} = \left(\sum x_i - n\right)\left(\frac{1}{a-1} - \frac{1}{a}\right) - \frac{n}{a} = 0 \;\Rightarrow\; \frac{\sum x_i}{n} = a.$$

The maximum likelihood estimate is $\boxed{a = \bar{x}}$.

9. If there are n students in the room then for the data 1, 3, 7 (occuring in any order) the likelihood is

$$p(\text{data}\,|\,n) = \begin{cases} 0 & \text{for } n < 7 \\ 1/\binom{n}{3} = \frac{3!}{n(n-1)(n-2)} & \text{for } n \geq 7 \end{cases}$$

Maximizing this does not require calculus. It clearly has a maximum when n is as small as possible. Answer: $\boxed{n = 7}$.

5 Bayesian updating: discrete prior, discrete likelihood

10. This is a Bayes' theorem problem. The likelihoods are

P(same sex | identical) = 1 P(different sex | identical) = 0
P(same sex | fraternal) = 1/2 P(different sex | fraternal) = 1/2

The data is 'the twins are the same sex'. We find the answer with an update table

hyp.	prior	likelihood	unnorm. post.	posterior
identical	1/3	1	1/3	1/2
fraternal	2/3	1/2	1/3	1/2
Tot.	1		2/3	1

So $\boxed{\text{P(identical} \mid \text{same sex)} = 1/2}$.

11. (a) The data is 5. Let H_n be the hypothesis the die is n-sided. Here is the update table.

hyp.	prior	likelihood	unnorm. post.	posterior
H_4	1	0	0	0
H_6	2	$(1/6)^2$	$2/36$	0.243457
H_8	10	$(1/8)^2$	$10/64$	0.684723
H_{12}	2	$(1/12)^2$	$2/144$	0.060864
H_{20}	1	$(1/20)^2$	$1/400$	0.010956
Tot.	16		0.22819	1

So $P(H_8|\text{data}) = 0.685$.

(b) We are asked for posterior predictive probabilities. Let x be the value of the next roll. We have to compute the total probability

$$p(x|\text{data}) = \sum p(x|H)p(H|\text{data}) = \sum \text{likelihood} \times \text{posterior}.$$

The sum is over all hypotheses. We can organize the calculation in a table where we multiply the posterior column by the appropriate likelihood column. The total posterior predictive probability is the sum of the product column.

hyp.	posterior to data	likelihood (i) $x = 5$	post. to (i)	likelihood (ii) $x = 15$	post. to (ii)
H_4	0	0	0	0	0
H_6	0.243457	1/6	0.04058	0	0
H_8	0.684723	1/8	0.08559	0	0
H_{12}	0.060864	1/12	0.00507	0	0
H_{20}	0.010956	1/20	0.00055	1/20	0.00055
Tot.	0.22819		0.13179		0.00055

So, (i) $p(x = 5|\text{data}) = 0.132$ and (ii) $p(x = 15|\text{data}) = 0.00055$.

12. (a) Solution to (a) is with part (b).

(b) Let θ be the probability of the selected coin landing on heads. Given θ, we know that the number of heads observed before the first tails, X, is a geo(θ) random variable. We have updating table:

Hyp.	Prior	Likelihood	Unnorm. Post.	Posterior
$\theta = 1/2$	1/2	$(1/2)^3(1/2)$	$1/2^5$	16/43
$\theta = 3/4$	1/2	$(3/4)^3(1/4)$	$3^4/2 \cdot 4^4$	27/43
Total	1	–	43/256	1

The prior odds for the fair coin are 1, the posterior odds are 16/27. The prior predictive probability of heads is $0.5 \cdot \frac{1}{2} + 0.75 \cdot \frac{1}{2}$. The posterior predictive probability of heads is $0.5 \cdot \frac{16}{43} + 0.75 \cdot \frac{27}{43}$.

6 Bayesian Updating: continuous prior, discrete likelihood

13. **(a)** $x_1 \sim \text{Bin}(10, \theta)$.

(b) We have prior:

$$f(\theta) = c_1\theta(1-\theta)$$

and likelihood:

$$p(x_1 = 6 \,|\, \theta) = c_2\theta^6(1-\theta)^4, \qquad \text{where} \quad c_2 = \binom{10}{6}.$$

The unnormalized posterior is $f(\theta)p(x_1|\theta) = c_1c_2\theta^7(1-\theta)^5$. So the normalized posterior is

$$f(\theta|x_1) = c_3\theta^7(1-\theta)^5$$

Since the posterior has the form of a beta$(8,6)$ distribution it must be a beta$(8,6)$ distribution. We can look up the normalizing coefficient $c_3 = \frac{13!}{7!\,5!}$.

(c) The 50% interval is

`[qbeta(0.25,8,6),  qbeta(0.75,8,6)] = [0.48330,  0.66319]`

The 90% interval is

`[qbeta(0.05,8,6),  qbeta(0.95,8,6)] = [0.35480,  0.77604]`

(d) If the majority prefer Bayes then $\theta > 0.5$. Since the 50% interval includes $\theta < 0.5$ and the 90% interval covers a lot of $\theta < 0.5$ we don't have a strong case that $\theta > 0.5$.

As a further test we compute $P(\theta < 0.5|x_1) =$ `pbeta(0.5,8,6) = 0.29053`. So there is still a 29% posterior probability that the majority prefers frequentist statistics.

(e) Let x_2 be the result of the second poll. We want $p(x_2 > 5|x_1)$. We can compute this using the law of total probability:

$$p(x_2 > 5|x_1) = \int_0^1 p(x_2 > 5|\theta)p(\theta|x_1)\, d\theta.$$

The two factors in the integral are:

$$\begin{aligned} p(x_2 > 5|\theta) &= \binom{10}{6}\theta^6(1-\theta)^4 + \binom{10}{7}\theta^7(1-\theta)^3 + \binom{10}{8}\theta^8(1-\theta)^2 \\ &\quad + \binom{10}{9}\theta^9(1-\theta)^1 + \binom{10}{10}\theta^{10}(1-\theta)^0 \\ p(\theta|x_1) &= \frac{13!}{7!5!}\theta^7(1-\theta)^5 \end{aligned}$$

This can be computed exactly or numerically in R using the `integrate()` function. The answer is $P(x_2 > 5\,|x_1 = 6) = 0.5521$.

7 Bayesian Updating: discrete prior, continuous likelihood

14. For a fixed θ the likelihood is

$$f(x|\theta) = \begin{cases} 1/\theta & \text{for } x \le \theta \\ 0 & \text{for } x \ge \theta \end{cases}$$

If Alice arrived 10 minutes late, we have table

Hypothesis	Prior	Likelihood for $x = 1/6$	Unnorm. Post	Posterior
$\theta = 1/4$	1/2	4	2	3/4
$\theta = 3/4$	1/2	4/3	2/3	1/4
Total	1	–	8/3	1

In this case the most likely value of θ is 1/4.

If Alice arrived 30 minutes late, we have table

Hypothesis	Prior	Likelihood for $x = 1/2$	Unnorm. Post	Posterior
$\theta = 1/4$	1/2	0	0	0
$\theta = 3/4$	1/2	4/3	2/3	1
Total	1	–	2/3	1

In this case the most likely value of θ is 3/4.

8 Bayesian Updating: continuous prior, continuous likelihood

15. (a) We have $\mu_{\text{prior}} = 9, \sigma^2_{\text{prior}} = 1$ and $\sigma^2 = 10^{-4}$. The normal-normal updating formulas are

$$a = \frac{1}{\sigma^2_{\text{prior}}} \qquad b = \frac{n}{\sigma^2}, \qquad \mu_{\text{post}} = \frac{a\mu_{\text{prior}} + b\bar{x}}{a+b}, \qquad \sigma^2_{\text{post}} = \frac{1}{a+b}.$$

So we compute $a = 1/1, b = 10000$, $\sigma^2_{post} = 1/(a+b) = 1/10001$ and

$$\mu_{\text{post}} = \frac{a\mu_{\text{prior}} + bx}{a+b} = \frac{100009}{10001} \approx 9.990$$

So we have posterior distribution $f(\theta|x = 10) \sim N(9.99990, 0.0099)$.

(b) We have $\sigma^2_{\text{prior}} = 1$ and $\sigma^2 = 10^{-4}$. The posterior variance of θ given observations $x_1, \ldots, x_n$ is given by

$$\frac{1}{\frac{1}{\sigma^2_{\text{prior}}} + \frac{n}{\sigma^2}} = \frac{1}{1 + n \cdot 10^4}$$

We wish to find n such that the above quantity is less than 10^{-6}. It is not hard to see that $n = 100$ is the smallest value such that this is true.

16. We have likelihood function

$$f(x_1, \ldots, x_5|\lambda) = \prod_{i=1}^{5} \lambda e^{-\lambda x_i} = \lambda^5 e^{-\lambda(x_1+x_2+\cdots+x_5)} = \lambda^5 e^{-2\lambda}$$

So our posterior density is proportional to:

$$f(\lambda)f(x_1, \ldots, x_5|\lambda) \propto \lambda^9 e^{-3\lambda}$$

The hint allows us to compute the normalizing factor. (Or we could recognize this as the pdf of a Gamma random variable with parameters 10 and 3. Thus, the density is

$$f(\lambda|x_1, \ldots, x_5) = \frac{3^{10}}{9!}\lambda^9 e^{-3\lambda}.$$

17. (a) Let X be a random decay distance.

$$Z(\lambda) = P(\text{detection} \,|\, \lambda) = P(1 \le X \le 20 \,|\, \lambda) = \int_1^{20} \lambda e^{-\lambda x}\, dx = \boxed{e^{-\lambda} - e^{-20\lambda}}.$$

(b) Fully specifying the likelihood (remember detection only occurs for $1 \le x \le 20$).

$$\text{likelihood} = f(x \,|\, \lambda, \text{detected}) = \frac{f(x \text{ and detected} \,|\, \lambda)}{f(\text{detected} \,|\, \lambda)} = \begin{cases} \frac{\lambda e^{-\lambda x}}{Z(\lambda)} & \text{for } 1 \le x \le 20 \\ 0 & \text{otherwise} \end{cases}$$

(c) Let Λ be the random variable for λ. Let $\overline{X} = 1/\Lambda$ be the random variable for the mean.

We are given that $\overline{X}$ is uniform on $[5, 30]$ $\Rightarrow$ $f_{\overline{X}}(\overline{x}) = 1/25$.

First we find $f_\Lambda(\lambda)$ by finding and then differentiating $F_\Lambda(\lambda)$.

$$F_\Lambda(\lambda) = P(\Lambda < \lambda) = P\left(\frac{1}{\Lambda} > \frac{1}{\lambda}\right) = P\left(\overline{X} > \frac{1}{\lambda}\right)$$

$$\begin{cases} 1 & \text{for } 5 < 1/\lambda \\ \frac{30-1/\lambda}{25} & \text{for } 5 < 1/\lambda < 30 \\ 0 & \text{for } 1/\lambda > 30 \end{cases} = \begin{cases} 1 & \text{for } \lambda < 5 \\ \frac{30}{25} - \frac{1}{25\lambda} & \text{for } 1/30 < \lambda < 1/5 \\ 0 & \text{for } \lambda < 1/30 \end{cases}$$

Taking the derivative we get

$$f_\Lambda(\lambda) = F'_\Lambda(\lambda) = \frac{1}{25\lambda^2} \text{ on } \frac{1}{30} < \lambda < \frac{1}{5}.$$

From part (b) the likelihood $f(x_i \,|\, \lambda) = \dfrac{\lambda e^{-\lambda x_i}}{Z(\lambda)}$. So the likelihood

$$f(\text{data} \,|\, \lambda) = \frac{\lambda^4 e^{-\lambda \sum x_i}}{Z(\lambda)^4} = \frac{\lambda^4 e^{-43\lambda}}{Z(\lambda)^4}$$

Now we have the prior and likelihood so we can do a Bayesian update:

Hypothesis	prior	likelihood	posterior
λ $(1/30 < \lambda < 1/5)$	$\dfrac{1}{25\lambda^2}$	$\dfrac{\lambda^4 e^{-43\lambda}}{Z(\lambda)^4}$	$c\,\dfrac{\lambda^2 e^{-43\lambda}}{Z(\lambda)^4}$

$$\text{Odds}\left(\frac{1}{\lambda} > 10\right) = \text{Odds}\left(\lambda < \frac{1}{10}\right) = \frac{P(\lambda < 1/10)}{P(\lambda > 1/10)}$$

$$= \frac{\int_{1/30}^{1/10} \text{posterior}\, d\lambda}{\int_{1/10}^{1/5} \text{posterior}\, d\lambda} = \frac{\int_{1/30}^{1/10} \frac{\lambda^2 e^{-43\lambda}}{Z(\lambda)^4}\, d\lambda}{\int_{1/10}^{1/5} \frac{\lambda^2 e^{-43\lambda}}{Z(\lambda)^4}\, d\lambda}$$

Using the R function `integrate()` we computed $\text{Odds}\left(\frac{1}{\lambda} > 10\right) \approx 10.1$.

9 NHST

18. (a) Our z-statistic is

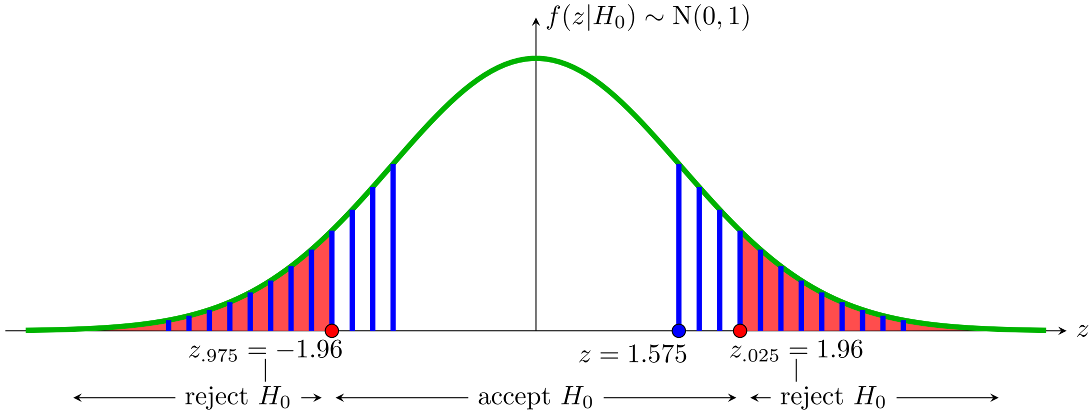
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{6.25 - 4}{10/7} = 1.575$$

Under the null hypothesis $z \sim \mathrm{N}(0, 1)$ The two-sided p-value is

$$p = 2 \times P(Z > 1.575) = 2 \times 0.0576 = 0.1152$$

The probability was computed from the z-table. We interpolated between $z = 1.57$ and $z = 1.58$ Because $p > \alpha$ we do not reject H_0.

(b) The null pdf is standard normal as shown. The red shaded area is over the rejection region. The area used to compute significance is shown in red. The area used to compute the p-value is shown with blue stripes. Note, the z-statistic outside the rejection region corresponds to the blue completely covering the red.



19. (a) Our t-statistic is

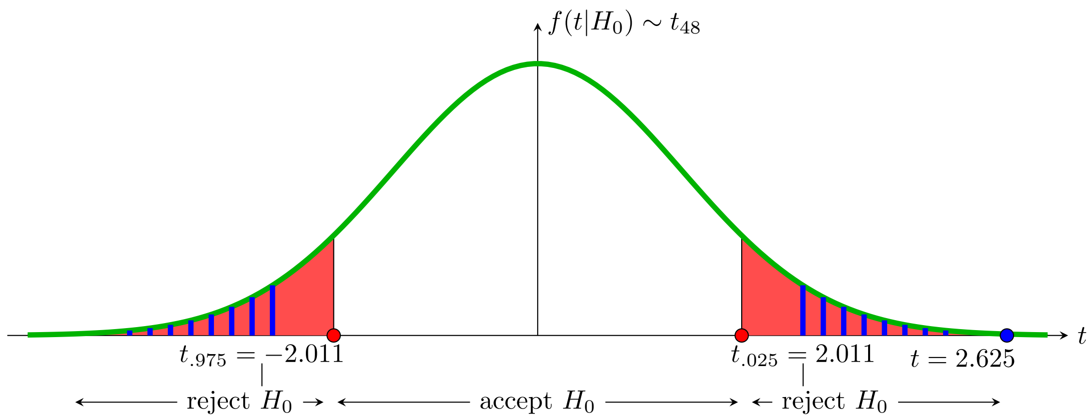
$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{6.25 - 4}{6/7} = 2.625$$

Under the null hypothesis $t \sim t_{48}$. Using the t-table we find the two-sided p-value is

$$p = 2 \times P(t > 2.625) < 2 \times 0.005 = 0.01$$

Because $p < \alpha$ we reject H_0.

(b) The null pdf is a t-distribution as shown. The rejection region is shown. The area used to compute significance is shown in red. The area used to compute the p-value is shown with blue stripes. Note, the t-statistic is inside the rejection region corresponds. This corresponds to the red completely covering the blue. The critical values for t_{48} we're looked up in the table.



20. Probability, MLE, goodness of fit

(a) (i) This is a binomial distribution. Let θ be the Bernoulli probability of success

$$p(x = k) = \binom{12}{k} \theta^k (1-\theta)^{1-k}, \text{ for } k = 0, 1, \ldots, 12.$$

(ii) This is a version of a geometric distribuiton

$$p(x = k) = \theta^k(1-\theta), \ \text{ for } k = 0, 1, \ldots$$

(b) (i) The likelihood function for n trials is

$$\begin{aligned} p(x_1, x_2, \ldots, x_n \,|\, \theta) &= \binom{12}{x_1}\theta^{x_1}(1-\theta)^{12-x_1}\binom{12}{x_2}\theta^{x_2}(1-\theta)^{12-x_2}\cdots\binom{12}{x_n}\theta^{x_n}(1-\theta)^{12-x_n} \\ &= c\theta^{\sum x_i}(1-\theta)^{\sum 12-x_i} \end{aligned}$$

To find the maximum we use the log likelihood. We also substitute $n\bar{x}$ for $\sum x_i$.

$$\ln(p(\text{data} \,|\, \theta) = \ln(c) + n\bar{x}\ln(\theta) + n(12-\bar{x})\ln(1-\theta).$$

To find the maximum we set the derivative to 0:

$$\frac{d\,\ln(p(\text{data} \,|\, \theta))}{d\,\theta} = \frac{n\bar{x}}{\theta} - \frac{n(12-\bar{x})}{1-\theta} = 0.$$

Solving for θ we get

$$\hat{\theta} = \frac{\bar{x}}{12}.$$

(ii) The likelihood function for n trials is

$$p(x_1, x_2, \ldots, x_n \,|\, \theta) = \theta^{x_1}(1-\theta)\theta^{x_2}(1-\theta)\cdots\theta^{x_n}(1-\theta) = \theta^{\sum x_i}(1-\theta)^n$$

To find the maximum we use the log likelihood. We also substitute $n\bar{x}$ for $\sum x_i$.

$$\ln(p(\text{data} \,|\, \theta)) = n\bar{x}\ln(\theta) + n\ln(1-\theta).$$

To find the maximum we set the derivative to 0:

$$\frac{d\,\ln(p(\text{data} \,|\, \theta))}{d\,\theta} = \frac{n\bar{x}}{\theta} - \frac{n}{1-\theta} = 0.$$

Solving for θ we get

$$\hat{\theta} = \frac{\bar{x}}{1+\bar{x}}.$$

(c) The sample mean is

$$\begin{aligned} \bar{x} &= \frac{\sum(\text{count} \times x)}{\sum \text{counts}} \\ &= \frac{18\cdot 0 + 12\cdot 1 + 7\cdot 2 + 10\cdot 3 + 3\cdot 4 + 2\cdot 5 + 3\cdot 6 + 2\cdot 7 + 1\cdot 8 + 1\cdot 9 + 0\cdot 10 + 1\cdot 11}{60} \\ &= 2.30 \end{aligned}$$

(d) Just plug $\bar{x} = 2.3$ into the formulas from part (b):
(i) $\hat{\theta} = \bar{x}/12 = 2.3/12 = 0.19167$ (ii) $\hat{\theta} = \bar{x}/(1+\bar{x}) = 2.3/3.3 = 0.69697$

(e) There were 60 trials in all. Using the the values for $\hat{\theta}$ in part (d) we have the following tables. The probabilities are computed using R, the expected values are just the probabilities time 60. The components of X^2 are computed using the formula

$$X_i^2 = (E_i - O_i)^2/E_i.$$

For Experiment 1:
H_0 is the data comes from a binomial(12, 0.193) distribution.
H_A is it comes from some other distribution.

x	0	1	2	3	≥ 4
$p(x)$	0.077819	0.221423	0.288763	0.228232	0.183762
Observed	18	12	7	10	13
Expected	4.6691	13.2854	17.3258	13.6939	11.0257
X_i^2	38.06090	0.12437	6.15395	0.99644	0.35351

The χ^2 statistic is $X^2 = \sum X_i^2 = 45.689$. There are 5 cells, so 3 degrees of freedom. The p-values is

$p =$ `1-pchisq(45.689,4) = 2.86e-9`

With this p-value we reject H_0 in favor of the explanation that the data comes from a different distribution.

For Experiment 2:
H_0 is the data comes from a geometric(0.698) distribution, You have to be careful here. We define geometric(θ) as the number of successes before the first failure, where θ is the probability of success. In R the geometeric distribution is the opposite. It gives the number of failures till the first success. So, for calculations in R we have we use `dgeom(x, 1 - 0.698)` etc.)

H_A is it comes from some other distribution.

x	0	1	2	3	≥ 4
$p(x)$	0.30303	0.21120	0.14720	0.10260	0.23597
Observed	18	12	7	10	13
Expected	18.1818	12.6722	8.8321	6.1557	14.1582
X_i^2	0.0018182	0.0356546	0.3800529	2.4007701	0.0947394

The χ^2 statistic is $X^2 = \sum X_i^2 = 2.9130$. There are 5 cells, so 4 degrees of freedom. The p-values is

$p =$ `1-pchisq(2.9130, 4) = 0.573`

With this p-value we do not reject H_0.

21. See the psets 7 and 8.