Worksheet 29 Solution

.5cm

1. I have a random variable with PDF xe^{-x} for $x \geq 0$ and 0 otherwise. Prove that this is a valid PDF, and give the mean, mode and variance. To verify it is a PDF, we note that the function is

   nonnegative and

   $$\int_0^\infty xe^{-x}dx = -xe^{-x}|_0^\infty + \int_0^\infty e^{-x} = 1$$

   Mean we take

   $$\int_0^\infty x^2e^{-x}dx = -x^2e^{-x}|_0^\infty + \int_0^\infty 2\int_0^\infty xe^{-x} = 2$$

   Variance we could do in the same way, or just realize that this is $\Gamma(4) - \Gamma(3)^2 = 2$.

   For the mode, we have $\frac{d}{dx}xe^{-x} = e^{-x} - xe^{-x}$. If we set this to 0, we see our distribution has 0 derivative at 1, so is maximized at 1.

2. You have a Poisson distribution, and take samples 4,5,9,3. What is the maximum likelihood estimate? No short cuts! What is a 95% confidence interval for this parameter?

   Our maximum likelihood estimate is maximizing the joint PDF

   $$\frac{e^{-\lambda}\lambda^4}{4!}\frac{e^{-\lambda}\lambda^5}{5!}\frac{e^{-\lambda}\lambda^9}{9!}\frac{e^{-\lambda}\lambda^3}{3!}$$

   We can equivalently find the maximum of the log of this

   $$-4\lambda + 21\log(\lambda) - \log(k)$$

   for some constant k. Taking the derivative and setting it to 0 we get

   $$-4 + \frac{21}{\lambda} = 0$$

   So we get the maximum likelihood λ as $\frac{21}{4}$, which you probably could have guessed.

   To find a 95% confidence interval we take 2 sample standard deviations around this maximum likelihood estimate. The standard deviation of the Poisson distribution is $\sqrt{\lambda} \approx 2.29$ so our answer is $(5.25 - \frac{2 \cdot 2.29}{2}, 5.25 + \frac{2 \cdot 2.29}{2})$

3. The Albany School of Witchcraft claims that their classes increase the "Magic Quotient" (MQ) of their students. Witches outside of the school have an average MQ of 10, with an unknown standard deviation. You test the MQ of 6 students and receive the following MQs

MQ
9
12
14
13
9
12

   Is this enough information to tell that ASW's classes work, with confidence level .05? Use a Z and T test and compare the answers. Which do you think is more accurate?

The sample mean is 11.5. The sample standard deviation for the Z test is 1.89. Therefore the corresponding Z statistic is $\frac{(\overline{X}-\mu_0)\sqrt{n}}{\sigma} \approx 1.94$. $0.5 - Z(1.94) \approx 0.0262$ so much smaller than α, the classes work.

For the T-test, we use the unbiased standard deviation $1.89 \cdot \sqrt{6/5} = 2.07$, so we multiply our previous statistic by $\sqrt{5/6}$, yielding 1.77. The T statistic for 5 degrees of freedom and significance level .05 is 2.02, meaning that we cannot conclude that the classes work!

4. A well-known study examined the number of children who sleep with different levels of ambient light in an attempt to explore whether nighttime ambient light affects vision. In a survey, researchers found that 9 of 172 children who slept in darkness developed myopia, and 31 of 232 who slept with a small night-light developed myopia. Construct a contingency table and test for an association between nighttime ambient light and myopia.

   We have 40 children developed myopia, and 232 who slept with a night light, so if these were independent we would expect $40 * 232/404$ to have myopia and sleep with a light. Through this method, the expected values we recover are 23, 209, 17, 155. Therefore our chi squared statistic is

$$\frac{(23-31)^2}{23} + \frac{(9-17)^2}{17} + \frac{(209-201)^2}{209} + \frac{(155-163)^2}{163} = 7.25$$

   There is one degree of freedom, and this is above the statistic 3.84, meaning we can not tell with certainty if there is independence.

5. $z = x^2 - xy - 2y^2$. Take $x = s + t$, $y = st$

   What are $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$?

   $z = (s+t)^2 - (s+t)st - 2s^2t^2$ so $\frac{\partial z}{\partial s}$ is $2s+2t-st-(s+t)t-4st^2$ and $\frac{\partial z}{\partial t}$ is $2s+2t-st-(s+t)s-4s^2t$.

6. The HAART data gives before and after measurements for 5 individuals. Calculate the best fit line and correlation coefficient.

Before	After
7.4	3.7
5.1	2.6
6.9	3.4
7.2	3.6
1.4	0.7

   For the correlation coefficient, we take $\frac{E(XY)-E(X)E(Y)}{\sigma_X \sigma_Y} = \frac{2.52}{2.52 * 1.25} = .80$. For the best fit line, we have $a = r\frac{\sigma_Y}{\sigma_X} = .4$. Our b value for the best fit line is $E(Y) - aE(X) = .58$

Worksheet 29 Solution

The minimum inhibitory concentration (MIC) of a new antibiotic is measured in 10 bacterial colonies. Two aditional, independent studies, are conducted and the MIC is measured in another 10 bacterial colonies in each of these studies. Suppose that the MIC is well described by a normal random variable with unknown mean and with standard deviation 0.043 $\mu g/mL$. What is the probability that at least one of the sample means from these three studies deviates by more that 0.01 $\mu g/mL$ from the true population mean? This is the compliment of if none of

them differ by more than 0.01. The probability of each of these differing by less than 0.01 is $2 \cdot Z(.01/.043) \approx 2 \cdot Z(.23) \approx .182$. This is so our answer is $1 - .182^3 \approx .994$