$P(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

1. A type of weeds grow in a farmer's field such that when he walks through his field he encounters, on average, 10 weeds every 50 yards. Suppose the weeds grow according to a Poisson process. Here comes the farmer now; he is about to walk in his field... $\lambda = 10$

   (a) Give the probability that he will walk 50 yards without encountering a single weed.

$P(X=0) = \frac{e^{-10}(10)^0}{0!} = \boxed{e^{-10}}$ ✓

   (b) Give the probability that he encounters at least one weed in the first 50 yards.

$P(x \geq 1) = 1 - P(x=0)$

$1 - \frac{e^{-10}(10)^0}{0!} = \boxed{1 - e^{-10}}$ ✓

   (c) Give an expression for the probability that he finds at least 10 weeds in the first 50 yards. You do not need to evaluate your expression.

$P(x \geq 10) = P(x=10) + P(x=11) + \ldots = 1 - P(x < 10) = 1 - P(x \leq 9) = \boxed{1 - \sum_{x=0}^{9} \frac{e^{-10}(10)^x}{x!}}$ ✓

   (d) Give the expected number of weeds he will encounter if he walks 300 yards in his field.

$P(Y < 300) = P(Y \leq 300) = \boxed{1 - e^{-6(10)}}$ ✓

$Y \sim \text{Exponential}(t \cdot \lambda) \longrightarrow Y \sim \text{Exponential}(6 \cdot 10)$

$\hookrightarrow$ 10 per 50, $\frac{300}{50} = \frac{30}{5} = 6$ sets of 50

$F(y) = 1 - e^{-x\lambda}$

   (e) Suppose he measures the distance from the first weed he encounters to the next weed he encounters. What probability distribution does this random variable have?

Exponential ✓

2. A certain purveyor's monster-size bags of potato chips have weights following a Normal distribution with mean 1.01 lbs and standard deviation 0.03. $X \sim \text{Normal}(1.01, 0.03^2)$

(a) Give the probability that a randomly sampled bag will weigh less than 0.99 lbs.

$P(X < 0.99) \rightarrow P(Z < -2/3)$

$\frac{0.99-1.01}{0.03} \rightarrow \frac{0.02}{0.03}$

$\boxed{P(X<0.99)=0.2546}$

(b) Give the probability that a randomly sampled bag will weight more than 0.99 lbs.

$\boxed{P(X>0.99) = 0.7454}$

$1-0.2546$

(c) Give the 97.72 percentile of bag weights.

$100-97.72 = 2.28$

$2.00 = \frac{x-1.01}{0.03}$

$0.06 = x - 1.01$

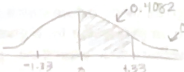
$\boxed{x = 1.07}$

(d) Suppose you take a random sample of 16 bags. Give the probability that the average weight of the 16 bags is less than 1 lb. $\bar{X}_{16}$

$P(\bar{X}_n < 1) \rightarrow P(Z<-1.33)$

$\boxed{P(\bar{X}_n < 1) = 0.0918}$

$\frac{1-1.01}{0.03/\sqrt{16}} = \frac{-0.01}{0.03/4} = \frac{-0.01}{0.0075} = -1.33$



(e) Suppose the bag weights were *not* normally distributed. What is required for the mean weights of samples of bags to have approximately a normal distribution?

sample size must be greater than or equal to 30 to apply the central limit theorem

3. A blood donation advocate is interested in the proportion of college freshmen who have, prior to starting college, donated blood. She draws a sample of 100 freshmen, among whom she finds that 15 donated blood before starting college.

(a) Give an expression for a 95% confidence interval for the true proportion of college freshmen who have donated blood prior to starting college. You do not need to evaluate the bounds of the interval.

$\hat{p}_n = \frac{15}{100}$

$\frac{15}{100} \pm z_{.025}\sqrt{\frac{\frac{15}{100}(1-\frac{15}{100})}{100}}$ where $z_{.025}$ is 1.96

$\alpha = .05$

$\frac{\alpha}{2} = .05/2 = .025$

(b) The researcher wishes to know whether the percentage of college freshmen who have given blood before starting college is less than 20%. Formulate a null and an alternate hypothesis which correspond to her research question.

$H_0: p \geq .20$

$H_1: p < .20$

(c) Compute the value of the test statistic (it is possible, with some arithmetic, do to this without a calculator) for testing your hypotheses in the previous part.

$z_{test} = \frac{.15 - .20}{\sqrt{\frac{.20(1-.20)}{100}}} = \frac{-.05}{\frac{\sqrt{.16}}{\sqrt{100}}} = \frac{-\frac{5}{100}}{\frac{40/100}{10}}$ $\quad \frac{40}{100}\cdot\frac{1}{10} = \frac{4}{100}$ $\quad \boxed{= -1.25}$ $\quad -\frac{5}{100}\cdot\frac{100}{4} = -\frac{5}{4}$

(d) Does the researcher reject H_0 at the $\alpha = 0.05$ significance level?

$-z_{\alpha=.05} = -1.6449$

thus, $-1.25 > -1.6449$

And thus fail to reject H_0

(e) Suppose the researcher plans to do a larger study; she wishes in the end to be able to construct a confidence interval for the true proportion of freshmen who have given blood before starting college with a margin of error no greater than 3 percentage points. Give an expression for a guess of the sample size needed for her to achieve this. You do not need to evaluate your expression.

$m = .03$

$n = \left(1.96\left(\frac{\sqrt{\frac{15}{100}(1-\frac{15}{100})}}{.03}\right)\right)^2$

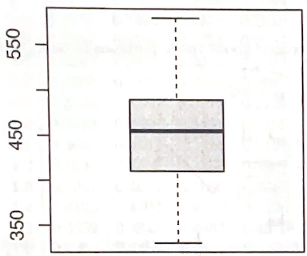
$z_{.05/2} = .025$

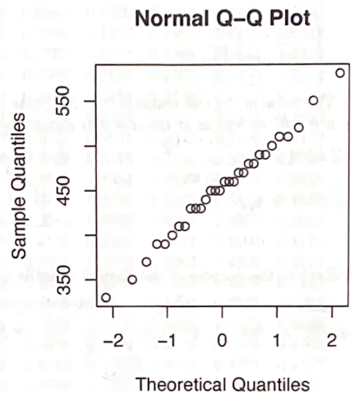
$= 1.96$

4. Municipal water quality investigators take 30 water samples and measure in each one the level of organic pollutants, obtaining the measurements below:

330	390	410	430	450	460	470	480	500	520
350	390	410	430	450	460	470	490	510	550
370	400	430	440	450	460	480	490	510	580

The average of all the measurements is 452 and the standard deviation is 56.47. Boxplots and a Normal quantile-quantile plot are shown.





(a) Give an expression for a 95% confidence interval for the true pollutant level in the water. You do not have to evaluate the endpoints of your interval.

$\bar{X}_n = 452$ $S_n = 56.47$ $t = 2.0452$

$452 \pm t_{29,.025} \frac{56.47}{\sqrt{30}}$

$$\boxed{\left(452 - 2.0452\left(\frac{56.47}{\sqrt{30}}\right), 452 + 2.0452\left(\frac{56.47}{\sqrt{30}}\right)\right)}$$

(b) Municipal authorities will take action if it is determined that the true pollutant level exceeds 430. Write down the relevant null and alternate hypotheses.

$H_0: \mu \leq 430$

$H_1: \mu > 430$

(c) Give an expression for the test statistic for testing the hypotheses in the previous part. You do not have to evaluate your expression.

$$\frac{\hat{X}_n - \mu_0}{\sigma/\sqrt{n}} \rightarrow \boxed{\frac{452-430}{56.47/\sqrt{30}}}$$

$\mu_0 = 430$ $n=30$

$\hat{X}_n = 452$

$\sigma = 56.47$

(d) The value of the test statistic is 2.134. State your decision about the hypotheses in part (b) at the $\alpha = 0.05$ as well as at the $\alpha = 0.01$ significance level.

$a = .05$ $n=30$

Reject H_0 if $T > t_{n-1,a}$

if $2.134 > 1.69$

Reject H_0 at $a = .05$

$a = .01$ $n=30$

Reject H_0 if $T > t_{n-1,a}$

if $2.134 > 2.46$

Do not Reject H_0 at $a = .01$

(e) Explain the purpose of the Normal quantile-quantile plot.

The normal QQ plot illustrates if a data set has normal distribution or not. Given that this normal QQ plot has a linear line of best fit, the data can be assumed to be normal,