Assignment 2

Karla Aniela Cepeda Zapata
D00242569

Hi Siobhán:

Please, find the solutions from Assignment 2, Section B.
The first three pages are my answer in handwriting (I tried to take the best pictures, and they seem better than last time), and also my answers in digital version.

Regards,
Karla Cepeda

Section B: Probability Distributions and Proportions

Section B: Probility distribution and proportions.

Question 1:

a. we have a number of trials $n = 18$ and a success of probability of $p = 0.25$ that the call is regarding billing.

- Given the information above, we just have two values: yes and no (i.e. 1 or 0, i.e. success or fail).
- Each call has a ~~relationship~~ historical probability of 25% of being call regarding billing.
- Calls unrelated from each other since it is stated that the sampling was random (i.e. independent, previous call did not affect probability of next calls).

∴ The probability distribution that we will use for this question ~~is the binomial distribution~~ is the binomial distribution

b. $n = 18$ $\quad p = 0.25$ $\quad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

$P(X=3) = \binom{18}{3}(0.25)^3(1-0.25)^{18-3}$

$P(X=3) = (816)(0.0156)(0.0133)$

$P(X=3) \approx 0.1704$

∴ The probability that three out of the 18 sample calls would be related to billing is 17.04%.

c. $n=18$ $\quad p = 0.25$ $\quad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

$P(X=0) = \binom{18}{0}(0.25)^0(1-0.25)^{18-0}$

$P(X=0) = (1)(1)(0.0056)$

$P(X=0) \approx 0.0056$

∴ The probability that zero calls out of 18 from sample will be related to billing is 0.56%.

d. $n=18$ $\quad p=0.25$ $\quad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

$P(X \leq 1) = P(X=0) + P(X=1)$

$P(X=1) = \binom{18}{1}(0.25)^1(1-0.25)^{18-1}$

$P(X=1) = (18)(0.25)(0.0075)$

$P(X=1) \approx 0.0338$

$P(X=0) \approx 0.0056$ (from previous question)

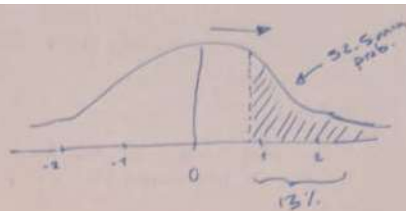
$P(X \leq 1) = 0.0056 + 0.0338$

$P(X \leq 1) \approx 0.0394$

∴ The probability that at most one call out of 18 sample calls will be related to billing is 3.94%

Question 2:

a. $x = 52.5 \quad \mu = 45 \quad \sigma = 2.5$



$P(X > 52.5)$

$z = \frac{X-\mu}{\sigma} = \frac{52.5-45}{2.5} \cong 3$

Looking at the z-score table:

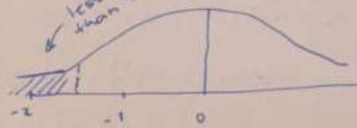
$P(z<3) \approx 0.99865$

but,

$1 - P(z<3) = 1 - 0.99865 \approx 0.0013.$

∴ The probability that the time require to fix a PC in more than 52.5 min is 0.13%.

b. $x = 41 \quad \mu = 45 \quad \sigma = 2.5$



$P(x<41)$

$z = \frac{x-\mu}{\sigma} = \frac{41-45}{2.5} = -1.6$

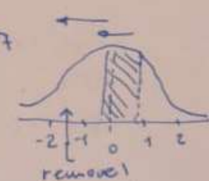
Looking at the z-score table:

$P(z<-1.6) =$

$P(z<-1.6) = 1 - P(z<1.6) \approx 1 - 0.94520 \approx 0.05480$

∴ The probability that the time require to fix a P.C. in less than 41 mins is 5.48%.

c. $P(x<41) \cong 0.0548, \quad P(X<52.5) \cong 0.9987$



$P(41<x<52.5) = P(x<52.5) - P(X<41)$

$= 0.9987 - 0.0548$

≈ 0.9439

∴ The probability that the time require to fix a PC between 41 and 52 min. is 94.39%.

d. $P(z<1.04) = 0.85$ ← looking at z-score table.

$z = 1.04$

$\mu = 45$

$\sigma = 2.5$

$z = \frac{X-\mu}{\sigma}$; we need X. ⇒ $X = (z * \sigma) + \mu$

$X = (1.04 * 2.5) + 45$

$X \approx 47.6$

∴ Roughly 47.6 min would be required at least to fix a P.C. within 85% of probability.

Question 3.

Baby breasted for a period of time VS mother's nationality.

Assumptions:

- Random sample ✓ Yes!.
- Independent observations in the cells. ✓ Yes!
- Expected value in each cell is greater than 5. ✓ Yes!

Let's work with the Chi Squared test.

Hypothesis:

H_0: Breasted period is independent of Nationality of mother.
H_1: Breasted period is dependent of Nationality of mother.

Degrees of freedom = (rows-1) X (columns-1) = 1

$$X_1^2 = \sum_{i=1}^{4} \frac{(\text{Observed}_i - \text{Expected}_i)^2}{\text{Expected}_i}$$

i	observed	Expected	Obs-Exp	$(obs-Exp)^2$	$\frac{(obs-Exp)^2}{Exp}$
1	60	54.88	5.12	26.2144	0.4776
2	10	15.12	-5.12	26.2144	1.7337
3	38	43.12	-5.12	26.2144	0.6079
4	17	11.88	5.12	26.2144	2.2206

$\Sigma = 5.0259$

$X_1^2 = 5.0259$

Looking at the Chi-squared table

Critical value$_{0.95}$ = 3.841

$P(X_1^2 < 3.841) = 0.95$, therefore critical region is > 3.841

Since $5.025 > 3.841$, we reject the null hypothesis.

∴ Breasted period is dependent on nationality of mother (i.e., there is a relationship between breasted period and nationality of mother).

Question 1:

1. Historical records show that 25% of all calls to customer service of a mobile phone company are billing related. A random sample of 18 calls was taken.

a. What probability distribution will you use to find probabilities and why?

By constructing a table to see the information given:

Proportion of calls received to customer service of a mobile phone company*		
	Yes	**No**
Billing	0.25	0.75

* Based on Historical records

We have a number of trials $\boldsymbol{n = 18}$ ***and a success probability of*** $\boldsymbol{p = 0.25}$ ***that the call is regarding billing.***

- ***Given the information above, we just have two values: yes or no (i.e. 1 or 0, i.e. success or fail).***
- ***Each call has a historical probability of 25% of being call regarding billing.***
- ***Calls unrelated from each other since it is stated that the sampling was random (i.e. independent, previous call did not affect probability of next calls).***

$\therefore \boldsymbol{The\ probability\ distribution\ that\ we\ will\ use\ for\ this\ question\ is\ the\ binomial\ distribution}.$

b. Find the probability exactly three will be billing related;

$$\boldsymbol{n = 18}$$

$$\boldsymbol{p_{call\ regarding\ billing} = 0.25}$$

$$\boldsymbol{P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}}$$

$$\boldsymbol{P(X = 3) = \binom{18}{3} (0.25)^3 (1-0.25)^{18-3}}$$

$$\boldsymbol{P(X = 3) = (816)(0.015625)(0.75)^{15}}$$

$$\boldsymbol{P(X = 3) = (816)(0.015625)(0.013363461)}$$

$$\boldsymbol{P(X = 3) \cong 0.1704}$$

$\therefore \boldsymbol{The\ probability\ that\ three\ out\ of\ the\ 18\ sample\ calls\ would\ be\ related\ to\ billing\ is\ 17.04}\%.$

c. Find the probability none will be billing related;

$$\boldsymbol{n = 18}$$

$$\boldsymbol{p_{call\ regarding\ billing} = 0.25}$$

$$\boldsymbol{P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}}$$

$$\boldsymbol{P(X = 0) = \binom{18}{0} (0.25)^0 (1-0.25)^{18-0}}$$

$$\boldsymbol{P(X = 0) = (1)(1)(0.0056)}$$

$$\boldsymbol{P(X = 0) \cong 0.0056}$$

$\therefore \boldsymbol{The\ probability\ that\ zero\ calls\ will\ be\ related\ to\ billing\ is\ 0.56}\%.$

d. Find the probability at most one will be billing related.

$$n = 18$$

$$p_{call\ regarding\ billing} = 0.25$$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

To find out the probability of at most one call will be regarding billing we need to sum up all probabilities of $P(X \leq 1)$.

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

Given that we have already had $P(X = 0) \cong 0.013363$, *the only probability which is missing is* $P(X = 1)$.

$$P(X = 1) = \binom{18}{1} (0.25)^1 (1 - 0.25)^{18-1}$$

$$P(X = 1) = (18)(0.25)(0.75)^{17}$$

$$P(X = 1) = (18)(0.25)(0.007516946)$$

$$P(X = 1) \cong 0.033826$$

Now that we have $P(X = 1)$, *let's add up all probabilities together:*

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$P(X \leq 1) = 0.0056 + 0.0338$$

$$P(X \leq 1) = 0.0394$$

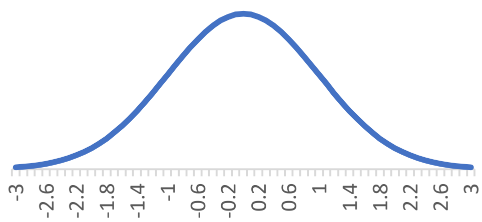
$\therefore$ ***The probability that at most one call out of the*** 18 ***sampled calls will be related to billing is*** 3.94%.

Question 2:

2. The amount of time it takes for a networking problem to be resolved is **normally distributed** with a mean of 45 minutes and a standard deviation 2.5 minutes. What is the probability that the time required to fix a particular P.C. picked at random will be…?

e. more than 52.5 minutes?

To find out the probability, let's use the standard normal distribution



$\boldsymbol{\mu = 0, \quad \sigma = 1, \quad z \sim N(0,1)}$

$\boldsymbol{X = 52.5}$

$\boldsymbol{\mu = 45}$

$\boldsymbol{\sigma = 2.5}$

$\boldsymbol{P(X > 52.5)}$

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{52.5 - 45}{2.5} = 3}$$

Looking at the Z-Score table

2.9	.99813	.99819
3.0	.99865	.99869
3.1	.99903	.99906
3.2	.99931	.99934
3.3	.99952	.99953
3.4	.99966	.99968
3.5	.99977	.99978

$\boldsymbol{P(z < 3) = 0.99865}$

However, we need above this probability since we have been asked to get greater than, so:

$\boldsymbol{P(z > 3) = 1 - P(z < 3) = 1 - 0.99865 = 0.0013}$

$\boldsymbol{\therefore \text{The probability that the time require to } fix \text{ a } P.C. \text{ in more than } 52.5 \text{ mins is } 0.13\%}$

f. less than 41 minutes?

$\boldsymbol{X = 41}$

$\boldsymbol{\mu = 45}$

$\boldsymbol{\sigma = 2.5}$

$\boldsymbol{P(X < 41)}$

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{41 - 45}{2.5} = -1.6}$$

Looking at the Z-Score table

1.4	.91924	.92073
1.5	.93319	.93448
1.6	$\boxed{.94520}$	.94630
1.7	.95543	.95637
1.8	.96407	.96485
1.9	.97128	.97193
2.0	.97725	.97778

Since we just have the positive (i.e. right) side of the z-score table, let's get the negative side by subtracting by one.

$P(z < -1.6) = 1 - P(z < 1.6) = 1 - 0.94520 = 0.05480$

$\therefore$ ***The probability that the time require to fix a P.C. in less than*** 41 ***mins is*** 5.48%

g. between 41 minutes and 52.5 minutes?

Given that we have already computed the probabilities of $P(X < 41)$ and $P(X < 52.5)$, then:

$P(X < 41) = 0.0548$

$P(X < 52.5) = 0.9987$

$P(41 < X < 52.5) = P(X < 52.5) - P(X < 41) = 0.9987 - 0.0548 = 0.9439$

$\therefore$ ***The probability that the time require to fix a P.C. between*** 41 ***and*** 52.5 ***mins is*** 94.39%

h. Determine the number of hours within which 85% of queries will be resolved.

Looking at the z-score table to find the z value:

STANDARD NORMAL DISTRIBUTION: Table Values Represent A

Z	.00	.01	.02	.03	.04	.05
0.0	.50000	.50399	.50798	.51197	.51595	.51994
0.1	.53983	.54380	.54776	.55172	.55567	.55962
0.2	.57926	.58317	.58706	.59095	.59483	.59871
0.3	.61791	.62172	.62552	.62930	.63307	.63683
0.4	.65542	.65910	.66276	.66640	.67003	.67364
0.5	.69146	.69497	.69847	.70194	.70540	.70884
0.6	.72575	.72907	.73237	.73565	.73891	.74215
0.7	.75804	.76115	.76424	.76730	.77035	.77337
0.8	.78814	.79103	.79389	.79673	.79955	.80234
0.9	.81594	.81859	.82121	.82381	.82639	.82894
1.0	.84134	.84375	.84614	.84849	$\boxed{.85083}$	.85314
1.1	.86433	.86650	.86864	.87076	.87286	.87493
1.2	.88493	.88686	.88877	.89065	.89251	.89435
1.3	.90320	.90490	.90658	.90824	.90988	.91149

$P(z < 1.04) = 0.85$

$z = 1.04$

$\mu = 45$

$\sigma = 2.5$

$$z = \frac{X - \mu}{\sigma}$$

$z \times \sigma = X - \mu$

$X = (z \times \sigma) + \mu$

$X = 1.04 \times 2.5 + 45$

$X = 47.6$

$\therefore$ ***Roughly, at least*** 47.6 ***mins would be required to fix a P.C. within*** 85% ***of probability.***

Question 3:

3. A survey was taken from a random sample of new mothers in a maternity hospital in Dublin to see how long they breastfed their babies for. The results are given in the table below. Test to see if there is a relationship between nationality and breastfeeding length.

		Baby breasted for		
		Up to 6 months	6 months or more	Total
Nationality of mother	Irish	60	10	70
	Non-Irish	38	17	55
	Total	98	27	125

Make sure in your answer to explain the null and alternative hypotheses, any assumptions needed and if they are met, results and interpretation of all the results. Conclude your findings.

Assumptions:

- ✓ *Random sample. Yes! In the sentences above said that the survey was taken randomly.*
- ✓ *Independent observations in the cells. Yes! There are no cells which overlap from each other. Additionally, it seems these groups are not related from each other.*
- ✓ *They are mutually exclusive.*
- ✓ $\boldsymbol{Expected\_value \geq 5}$*. Yes! As you can see in the table below, the values of the column* $\boldsymbol{Expected\_value}$ *are greater than five.*

Let's work with the Chi Squared Test.

Formulating the Null Hypothesis and Alternative Hypothesis:

- $\boldsymbol{H_0}$ = *Breasted period is independent of Nationality of Mother (i.e. There is no relationship between Breasted period and Nationality of mother).*
- $\boldsymbol{H_1}$ =*Breasted period is dependent of Nationality of Mother (i.e. There is relationship between Breasted period and Nationality of mother).*

i	$Observed$	$Expected$	$(Observed - Expected)$	$(Observed - Expected)^2$	$\frac{(Observed - Expected)^2}{Expected}$
1	60	54.88	5.12	26.2144	0.477668
2	10	15.12	-5.12	26.2144	1.733757
3	38	43.12	-5.12	26.2144	0.607941
4	17	11.88	5.12	26.2144	2.206599
					$\sum$ =**5.025964**

$$\boldsymbol{Degree\ of\ freedom = (num_{rows} - 1) \times (num_{columns} - 1) = (2 - 1) \times (2 - 1) = 1}$$

$$\chi_1^2 = \sum_{i=1}^{4} \frac{(\boldsymbol{Observed_i - Expected_i})^2}{\boldsymbol{Expected_i}} = \boldsymbol{5.024}$$

$\boldsymbol{With\ significant\ level\ \alpha = 0.05, and\ one\ tailed\ since\ Chi -}$
$\boldsymbol{Square\ Distribution\ is\ squared}$.

Degrees of Freedom	Values of P									
	0.005	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990	0.995
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.01	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860

$\boldsymbol{Critical\_value_{0.95} = 3.841}$

$\boldsymbol{P(\chi_1^2 < 3.841) = 0.95}$

Therefore, critical region is: $\boldsymbol{\chi_1^2 > 3.841}$

Since $\mathbf{5.025 > 3.841}$*, we reject the null hypothesis:*

$\therefore$ *Breasted period is dependent on nationality of mother (i.e. there is a relationship between breasted period and nationality of mother).*