Practice Problems Solution

Problems on Probability Solution

Solution Q.1

a) Pr(Satisfied) = 240/650 = 36.9%

b) Pr(Risk of Attrition) = Pr(Dissatisfied or Highly Dissatisfied) = Pr(Dissatisfied) + Pr(Highly Dissatisfied) = 37.7%

c) Pr(An employee will attrite) = Pr(Employee dissatisfied AND Employee attrite) = Pr(Employee dissatisfied) * Pr(Employee attrite|Dissatisfied) = 0.377 * 0.4 = 0.1508 = 15%
This result is through application of conditional probability results. By definition Pr(A∩B) = Pr(A)*Pr(B|A)

d) Pr(At least High School) = Pr(High school) + Pr(Some college) + Pr(College Grad) + Pr(Post-Grad) = 1 – Pr(Did not complete school) = 1 – 100/650 = 84.6%

e) Pr(Satisfied | College Grad or Post Grad) = 180 / 275 = 65.4%

f) Pr(Not dissatisfied | at least some college) = 325 / 425 = 76.5%

g) Pr(College Grad or Post-grad | Not satisfied) = 95 / 410 = 23.2%

h) Pr(Not highly dissatisfied | No college education) = 135 / 225 = 60%

Solution Q.2

a. Pr(Waiting for more than 5 minutes) = Pr(Waiting time > 5) = 0.15

b. Pr(Customers do not wait) = Pr(X = 0) = 0.2

c. Expected waiting time = E(X) = $\sum x \Pr[X = x]$ = 2.71

Waiting time (in minutes)	Probability	x Pr(x)

0	0.2	0
1	0.18	0.18
2	0.16	0.32
3	0.12	0.36
4	0.1	0.4
5	0.09	0.45
6	0.08	0.48
7	0.04	0.28
8	0.03	0.24

Problems on Binomial and Normal Distributions Solution

<u>Solution Q.1:</u>

This is a binomial probability problem.
No of trials = n = 15.
Each trial can result in success = not meeting pollution requirement.
Pr(success) = π= 0.6.
Define X = No. of successes in 15 trials = No of automobiles not meeting pollution requirement.

a) Probability that all 15 fails the inspection = Pr(X = 15) = 0.05%

```
> dbinom(15, 15, 0.6)
[1] 0.000470185
```

b) Pr(X = 8) = 17.7%

```
> dbinom(8, 15, 0.6)
[1] 0.1770837
```

c) Pr(Seven or less passes inspection) = $Pr(X \geq 8) = 1 - Pr(X < 8) = 1 - Pr(X \leq 7)$
Note that in this problem success is defined as NOT passing.

```
> 1 - pbinom(7, 15, 0.6)
[1] 0.7868968
```

We can also define success as passing the inspection. Then Pr(success) = 0.4. If Y = No of automobiles passing inspection, then we require $Pr(Y \leq 7)$

```
> pbinom(7, 15, 0.4)
[1] 0.7868968
```

d) $E(X) = n\pi = 15*0.6 = 9$ $Var(X) = n\pi(1 - \pi) = 15*0.6*0.4 = 3.6$

Expected value or mean of a binomial distribution is the number of successes you expect to see among n trials. Variance is a measure of the spread of the distribution.

Solution Q.2:

This is a binomial probability problem. For different parts success probability changes but number of trials stays the same.
No of trials = n = 10.

a) X = No rated outstanding. Pr(success) = 0.1. Pr(X = 2) = 19.4%

```
> dbinom(2, 10, 0.1)
[1] 0.1937102
```

b) X = No rated outstanding. Pr(success) = 0.1. Pr(X ≥ 2) = 1 – Pr(X ≤ 1) = 26.4%

```
> 1 - pbinom(1, 10, 0.1)
[1] 0.2639011
```

c) Y = No rated outstanding or excellent. Pr(success) = 0.1 + 0.75 = 0.85
Pr(Y = 8) = 27.6%

```
> dbinom(8, 10, 0.85)
[1] 0.2758967
```

d)W = No rated unsatisfactory. Pr(success) = 0.05. Pr(W = 0) = 59.9%

```
> dbinom(0, 10, 0.05)
[1] 0.5987369
```

Solution Q.3:

This is a binomial problem with n = 5 and success defined as admission. Pr(success) = 0.7, as given.
If success probability is 0.7 then probability that only 1 out of 5 is admitted, i.e. Pr(X=1) = 3%

```
> dbinom(1, 5, 0.7)
[1] 0.02835
```

Since this probability is very small, though not impossible, this event is not likely to occur. Hence our conclusion that in this instance, probability of acceptance does not seem to be 70%, but lower.

Note that this is a heuristic (informal) argument. When we go for statistical inference, such problems will be taken up in a more formal manner.

<u>Solution Q.4:</u>

This is a problem of normal probability distribution. Though the distribution is not mentioned, in absence of any other information we assume normality in the population.

X = cell phone bill amount. $X \sim N(\mu = 850, \sigma = 150)$

a) $Pr(X > 1200) = 0.98\%$

```
> 1 - pnorm(1200, 850, 150)
[1] 0.009815329
```

b) $Pr(750 \leq X \leq 1200) = 73.8\%$

```
> pnorm(1200, 850, 150) - pnorm(750, 850, 150)
[1] 0.7376921
```

c) $Pr(X \leq 650) = 9\%$

```
> pnorm(650, 850, 150)
[1] 0.09121122
```

d) Let the amount be M. $Pr(X \geq M) = 15\% \Rightarrow 1 - Pr(X < M) = 0.85 \Rightarrow M = 1005.46$

```
> qnorm(0.85, 850, 150)
[1] 1005.465
```

e) Let the amount be L. $Pr(X \leq L) = 0.25 \Rightarrow M = 748.83$

```
> qnorm(0.25, 850, 150)
[1] 748.8265
```

<u>Solution Q.5:</u>

Two different normal populations can only be compared if they are standardized. A normal distribution $N(\mu, \sigma)$ can be standardized as follows:

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

Z ~ N(0, 1) distribution. Z is known as a standard normal variable.
In this problem Z-scores are calculated from two different distributions by standardizing as above.
Z-score from Company A: Z1 = (50 – 40)/15 = 0.667
Z-score from Company B: Z2 = (50 – 45)/10 = 0.5

A higher score on Z scale is indicative of a better performance. Hence Ms. Z has done better in the test for Company A

Problems on Central Limit Theorem Solutions

<u>**Solution Q.1:**</u>

Trick here is to identify when to apply CLT.

a) X ~ N(5000, 500)
   Pr(X ≥ 5250) = 1 – Pr(X < 5250) = 30.8%

```
> 1 - pnorm(5250, 5000, 500)
[1] 0.3085375
```

b) Here interest is in the distribution of sample mean. Sample size n = 25
   X ~ N(5000, 500). Hence $\bar{X}$ ~ N(5000, 500/$\sqrt{25}$ = 100)

   Pr($\bar{X}$ ≥ 5250) = 1 – Pr($\bar{X}$ < 5250) = 0.6%

```
> 1 - pnorm(5250, 5000, 100)
[1] 0.006209665
```

<u>**Solution Q.2:**</u>

Within US$10,000 of population mean implies that sample mean $\bar{X}$ will be between [μ - 10000, μ + 10000]

a) Sample size n = 40 $\bar{X}$ ~ N(168000, 40000/$\sqrt{40}$ = 6324.55)
   Pr[168000-10000 ≤ $\bar{X}$ ≤ 168000+10000] = 88.6%

```
> pnorm(178000, 168000, 6324.55) - pnorm(158000, 168000, 6324.55)
[1] 0.886154
```

b) Sample size n = 40 $\bar{X}$ ~ N(117000, 25000/$\sqrt{40}$ = 3952.85)
   Pr[117000-10000 ≤ $\bar{X}$ ≤ 117000+10000] = 98.8%

```
> pnorm(127000, 117000, 3952.85) - pnorm(107000, 117000, 3952.85)
[1] 0.9885879
```

c) Among the female population we have a higher probability of getting a sample estimate with 10,000 of the true population mean. This is because for the females, population standard deviation is smaller.

d) Sample size n = 100 $\bar{X} \sim N(168000, 40000/\sqrt{100} = 4000)$

$Pr[\bar{X} \leq 164000] =$ 88.6%

```
> pnorm(164000, 168000, 4000)
[1] 0.1586553
```