Supplementary exercise 5.49 of IPS7e

Let X = number of word errors caught by the proofreader out of the 20 in the essay, and let $Y = 20 - X$ be the number of word errors missed.

(a) The distribution for X should be binomial $B(20, 0.7)$, and the distribution for Y should be binomial $B(20, 0.3)$ because the probability of missing a word is $1 - 0.7 = 0.3$. We can write $X \sim B(20, 0.7)$ and $Y \sim B(20, 0.3)$.

To facilitate the understanding of the following, we display the probability functions of both distributions using Minitab (where pcaught corresponds to the distribution for X and pmissed to the distribution for Y).

```
MTB > Name c1 "x"
MTB > Set 'x'
DATA>   1( 0 : 20 / 1 )1
DATA>   End.
MTB > Name c2 "pcaught"
MTB > PDF 'x' 'pcaught';
SUBC>   Binomial 20 .7.
MTB > Name c3 "pmissed"
MTB > PDF 'x' 'pmissed';
SUBC>   Binomial 20 .3.
MTB > Print 'x' 'pcaught' 'pmissed'.
```

Data Display

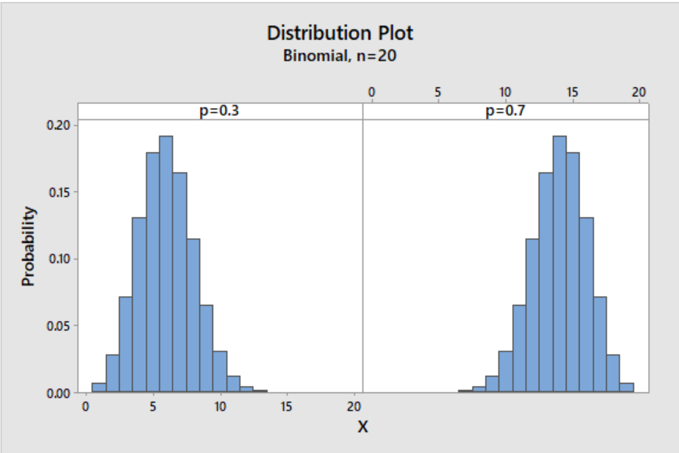
Data

Row	x	pcaught	pmissed
1	0	0.000000	0.000798
2	1	0.000000	0.006839
3	2	0.000000	0.027846
4	3	0.000001	0.071604
5	4	0.000005	0.130421
6	5	0.000037	0.178863
7	6	0.000218	0.191639
8	7	0.001018	0.164262
9	8	0.003859	0.114397
10	9	0.012007	0.065370
11	10	0.030817	0.030817
12	11	0.065370	0.012007
13	12	0.114397	0.003859
14	13	0.164262	0.001018
15	14	0.191639	0.000218
16	15	0.178863	0.000037
17	16	0.130421	0.000005
18	17	0.071604	0.000001
19	18	0.027846	0.000000
20	19	0.006839	0.000000
21	20	0.000798	0.000000

We also include a display using the Graph-Probability Distribution Plot menu, without any direct calculations in the worksheet. You can get both distributions in the same display by choosing the submenu "Vary Parameters".

```
MTB > DPlot;
SUBC>   Distribution;
SUBC>     Binomial 20 0.3;
SUBC>   Distribution;
```

```
SUBC>     Binomial 20 0.7;
SUBC>   Panel;
SUBC>   Same 2 1.
```



(b) The event "missing 9 or more words" is expressed as $Y \geq 9$. The Minitab listing above allows us to calculate

$$P(Y \geq 9) = 0.06537 + 0.03082 + 0.01201 + 0.00386 + 0.00102 + 0.00022 + 0.00004 = 0.1133.$$

Alternatively, Table 1 from Stephens gives

$$P(Y \geq 9) = 0.065 + 0.031 + 0.012 + 0.004 + 0.001 = 0.113.$$

Another possibility would be to use a cumulative probability calculation. For this, we calculate $P(Y \geq 9) = 1 - P(Y \leq 8) = 1 - 0.8867 = 0.1133$ or directly $P(Y \geq 9) = P(X \leq 11) = 0.1133$. We finally show Minitab code and outputs for these calculations:

```
MTB > CDF 8;
SUBC>   Binomial 20 0.3.
MTB > CDF 11;
SUBC>   Binomial 20 0.7.
```

Cumulative Distribution Function

Binomial with n = 20 and p = 0.3

x	P(X ≤ x)
8	0.886669

Cumulative Distribution Function

Binomial with n = 20 and p = 0.7

x	P(X ≤ x)
11	0.113331