STAT 35, Interactive and Computational Probability
UCLA Statistics

http://www.stat.ucla.edu/~dinov/courses_students.html

SOLUTIONS TO HOMEWORK 4

Solutions

Question 4_1

(a) $E(X) = \Sigma x P(x)$
$= 0(0.08) + 1(0.15) + 2(0.45) + 3(0.27) + 4(0.05)$
$= \mathbf{2.06}$

(b) $V(X) = \Sigma(x-\mu)^2 P(x)$
$= (0 - 2.06)^2(0.08) + (1 - 2.06)^2(0.15) + (2 - 2.06)^2(0.45)$
$+ (3 - 2.06)^2(0.27) + (4 - 2.06)^2(0.05)$
$= 0.3395 + 0.1685 + 0.0016 + 0.22386 + 0.1882$
$= \mathbf{0.9364}$

(c) $sd(X) = \sqrt{V(X)}$
$= 0.9364^{1/2}$
$= \mathbf{0.9677}$

(d) $V(X) = E(X^2) - \mu^2$
$= [0^2(0.08) + 1^2(0.15) + 2^2(0.45) + 3^2(0.27) + 4^2(0.05)] – (2.06)^2$
$= 5.18 - 4.2436$
$= \mathbf{0.9364}$

And we see that the answers from parts (b) and (d) are the same, which confirms that both methods of calculating the variance are equivalent.

Question 4_2

(a) $E(X) = 13.5\ (0.2) + 15.9\ (0.5) + 19.1\ (0.3) = \mathbf{16.38}$
$E(X^2) = 13.5^2(0.2) + 15.9^2(0.5) + 19.1^2(0.3) = \mathbf{272.3}$
$E(X^3) = 13.5^3(0.2) + 15.9^3(0.5) + 19.1^3(0.3) = \mathbf{4592}$
$V(X) = E(X^2) – (E(X))^2 = 272.298 - 268.3044 = \mathbf{3.994}$

(b) $E(25X - 8.5) = 25E(X) - 8.5 = 25(16.38) - 8.5 = \mathbf{\$401}$

(c) $Var(aX + b) = a^2 Var(X) = 25^2(3.9936) = 625(3.9936) = \mathbf{2496}$

Notice that the variance does not depend on b, which in this case is -8.5. Variance is a measure of the spread (variability) of the prices paid by different consumers for this item.

(d) $h(X) = X - 0.01X^2$

$E(h(X)) = E(X) - 0.01E(X^2) = 16.38 - 0.01(272.298) = \mathbf{13.68}$

Here we have used the rule that the expected value of a linear combination is the linear combination of the expected values.

Question 4_3

The number of heads should follow the Binomial distribution, with the given values of n and p. For this distribution, the moments are calculated as follows:

Theoretical Mean $= np$
$= (14)(0.3) = \mathbf{4.2}$

Theoretical Std. Deviation $= \sqrt{(np(1-p))}$
$= \sqrt{((14)(0.3)(0.7))} = \mathbf{1.715}$

The moments of the sample will differ from these values because the sample is relatively small. The standard error of the sample mean is $\sigma/\sqrt{n}$ i.e. $1.715/\sqrt{n}$. Thus the bigger the sample is, the more likely the sample mean is to be within 0.01 of the theoretical value. Note that since the standard error decreases in inverse proportion to $\sqrt{n}$, it decreases slowly as n increases. A large sample will be needed. We can show likewise for the sample standard deviation.

Even when one reaches a point where the sample moments are within 0.01 of the population moments, they are not guaranteed to stay this way if 10 more observations are made. This is because the observations, and these statistics that are based on them, are random.

Question 4_4

Just by looking at a graph of the distribution, one sees that nearly all of the probability is concentrated on values less than 35. Thus it is evident that the one almost never needs to clip 35 or more coupons to get 20 distinct ones. Looking at the table of the probability function, it shows that the exact value of the probability is 0.00153 + 0.00077 = 0.0023.

As for the probability of clipping at most 24 coupons to get 20 distinct ones, we may calculate the theoretical value again from the table of probabilities given:

$$\begin{aligned} P(W \leq 24) &= P(W = 20) + P(W = 21) + P(W = 22) + P(W = 23) + P(W = 24) \\ &= 0.01202 + 0.04569 + 0.09275 + 0.13368 + 0.15348 \\ &= \mathbf{0.43762} \end{aligned}$$

There may be some discrepancy between this theoretical value and the empirical probability observed in an experiment, but this difference should shrink as more and more observations are counted in the experiment.