**Exercise 5.4: Problem 2**

**(a).**

We can know that choosing the observations from original sample is a classical model of probability, so the probability of the first bootstrap observation is the *j*th observation = 1/n, so the probability that the first bootstrap observation is not the *j*th observation is 1-(1/n) = (n-1)/n.

**(b).**

The probability that the first bootstrap observation is not the *j*th observation is also (n-1)/n.

**(c).**

Since each choosing is independent, so the joint probability = ((n-1)/n)^n.

**(d).**

P(in)=1-P(out)=1-(4/5)^5=0.67

**(e).**

P(in)=1-P(out)=1-(99/100)^100=0.63

**(f).**

P(in)=1-P(out)=1-(9999/10000)^10000=0.63

**Exercise 6.8: Problem 1**

**(a).**

The model from best subset has the smallest training RSS, since model form best subset method choose *k* predictors independently, it need not contain predictors form *k-1* model.

**(b).**

We do not know; it depends on the method of validation and the training and test set.

**(c).**

**i.** True. According to the definition of forward stepwise.

**ii.** True. According to the definition of backward stepwise.

**iii.** False. Two methods are independent. For example, the forward method may choose x1, x3 in the k model (k=2), but the backward method may choose x1, x2 and x4 in k+1 model.

**iv.** False, two methods are independent.

**v.** False. Every step in best subset method is independent.