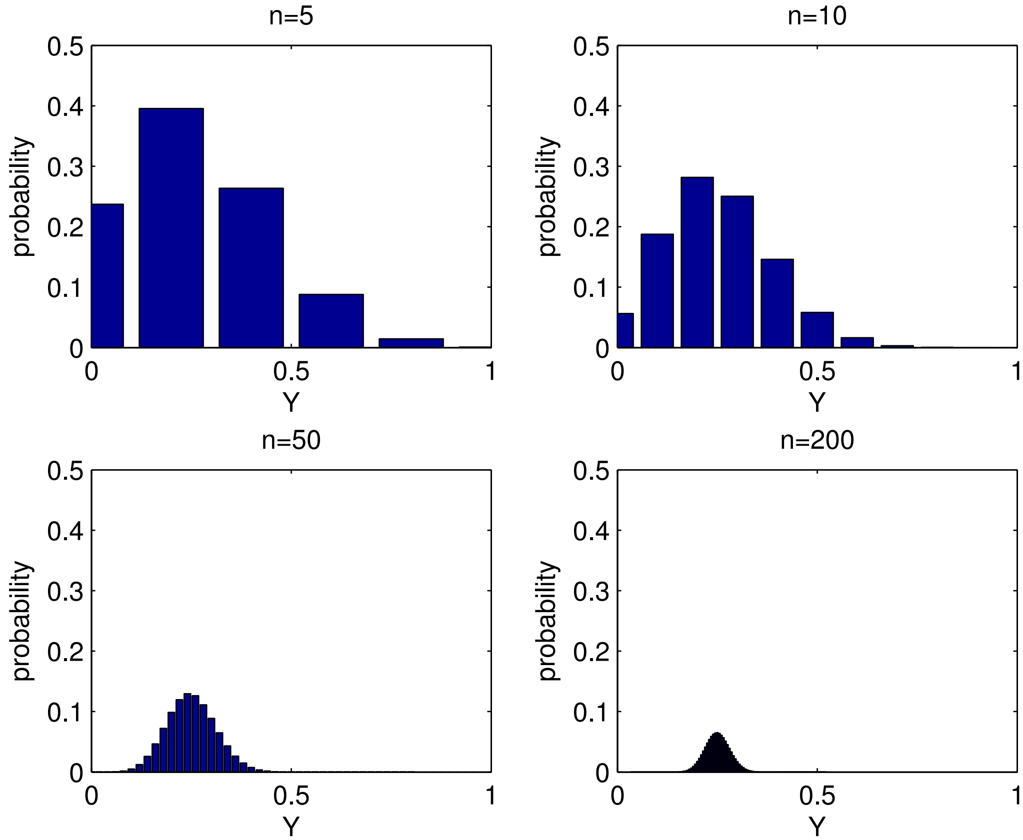
University of Illinois, Urbana–Champaign Fall 2007

ECE 313: Solutions to Homework 5

28. (a) $nY = \sum_{i=0}^{n} X_i \sim \mathsf{Binom}(n, 0.25)$.

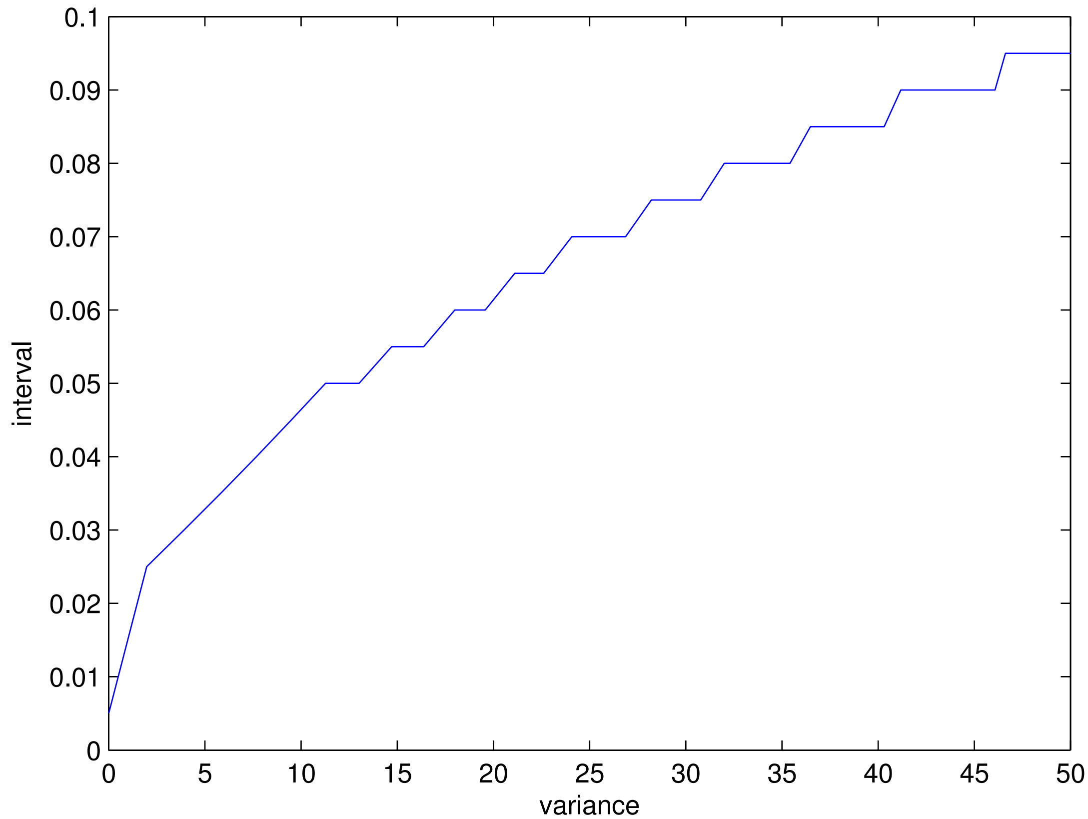


(b) Using Matlab to compute the interval that gives 99% confidence level. $\mathsf{E}(Y) = p = 0.25$. Here is the Matlab code:

```
for n = [5 10 50 200]
    e = round(0.25*n);
    p = binopdf(e, n, 0.25);
    i = 1;
    while p < 0.99
        p = p+ binopdf(e-i, n, 0.25);
        p = p+ binopdf(e+i, n, 0.25);
        i = i+1;
    end
    prob = p            % accumulated probability in the interval
    intval = i/n    % uni-lateral interval size
end
```

For $n = 5$: interval $\pm 0.8, p = 0.9990$. $n = 10$: interval $\pm 0.4000, p = 0.9965$. $n = 50$: interval $\pm 0.1800, p = 0.9953$. $n = 200$: interval $\pm 0.0850, p = 0.9931$. As n increases, the size of interval decreases. When $n \to \infty$, the interval approaches to 0. Intuitively, the average of value X for a large number of experiments approaches to the probability that X happens.

(c) Let p increases in steps of 0.01. Plot the result. When the probability of Heads(Tails) outweighs the probability of Tails(Heads), the variance of the average decreases as we increase in confidence that most of coins will get Heads(Tails). More specifically, when $p = 1$, the interval decreases to the 0 (or the minimal interval step in discrete case) and so does the trend of $p = 0$. When p is close 0.5, the variance peaks and it becomes most difficult for the frequency of the experiment to get close to the probability of one coin.



29. (a) Axiom 1: $Q(A) = P(A|B) = \frac{P(AB)}{P(B)}$. Since $P(AB) \geq 0, P(B) \geq 0, P(AB) \leq P(B)$, then $0 \leq Q(A) \leq 1$.

Axiom 2: $Q(\Omega) = Q(A) + Q(A^c) = \frac{P(AB)+P(A^cB)}{P(B)} = 1$

Axiom 3: For any sequence of mutually exclusive events $A_1, A_2, \ldots$, given that $P(\cdot)$ is a probability measure that satisfies Axiom 3, we have,

$$Q(\cup_{i=1}^{\infty} A_i) = \frac{P(\cup_{i=1}^{\infty} A_i B)}{P(B)} = \frac{\sum_{i=1}^{\infty} P(A_i B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B) = \sum_{i=1}^{\infty} Q(A_i)$$

(b) Since $Q(A)$ is a probability measure, we could define conditional probability on Q. So $Q(A|C) = P(A|BC), Q(C) = P(C|B), Q(C|A) = P(C|AB), Q(A) = P(A|B)$. Then we have,

$$P(A|B,C)P(C|B) = Q(A|C)Q(C) = Q(AC) = Q(C|A)Q(A) = P(C|AB)P(A|B)$$

Hence, include also as alternative derivation:

$$P(A|B,C) \equiv P(A|BC) = \frac{P(ABC)}{P(BC)} = \frac{P(C|AB)P(AB)}{P(C|B)P(B)}$$
$$= \frac{P(C|AB)P(A|B)P(B)}{P(C|B)P(B)} = \frac{P(C|AB)P(A|B)}{P(C|B)} \equiv \frac{P(C|A,B)P(A|B)}{P(C|B)}$$

30. Let R be the event that the blood result is positive. Let D be the event that a person carries the disease. We know that $P(R|D) = 0.95, P(R|D^c) = 0.01, P(D) = 0.005$. We want to find

$$P(D|R) = \frac{P(DR)}{P(R)} = \frac{P(R|D)P(D)}{P(DR) + P(D^cR)} = \frac{P(R|D)P(D)}{P(R|D)P(D) + P(R|D^c)P(D^c)} = 0.3231$$

Even though the blood test result is sensitive, the chance of getting false positive result weighs over because the disease is very rare in population. So the result is reasonable.

31. (a) The error rate distribution for the whole document is unknown.
    (b) According to Axiom 1, the smallest probability is zero.
    (c) Yes. Let C_i be the event that ith letter is typed correctly. We know that $P(C_2^c|C_1^c) = q_1, P(C_2^c|C_1) = q_2$. Then we have $P(C_1C_2) = P(C_2|C_1)P(C_1) = (1 - P(C_2^c|C_1))(1 - P(C_1^c)) = (1 - q_2)(1 - p) = 1 - p - q_2 + pq_2$.
    (d) On the other hand, we have $P(C_1C_2) = P(C_2) - P(C_1^cC_2) = 1 - P(C_2^c) - P(C_2|C_1^c)P(C_1^c) = 1 - p - (1 - q_1)p = 1 - 2p + q_1p$. Then, two equation should equal that $1 - p - q_2 + pq_2 = 1 - 2p + q_1p$. And they should comply probability law such that $0 \leq 1 - 2p + q_1 = 1 - p - q_2 + pq_2 \leq 1 - p$ and $0 \leq p_1, p_2 \leq 1$. Solving for it we get $q_1 = q_2\frac{p-1}{p} + 1$, where $0 \leq q_2 \leq \min\{1, \frac{p}{1-p}\}$.
    (e) Under independence assumption, $P(C_1C_2) = (1 - p)^2$. When $(1 - p)^2 = q - 2p + q_1p$, it renders the answer $q_1 = q_2 = p$ for (i). Intuitively, the answer is reasonable because the probability that the second letter is typed wrongly shouldn't depend on the condition of the first one. Similarly, we can get $q_2 < p$ for (ii) and $q_2 > p$ for (iii) in addition to the constraints we obtained in (d).

32. Let M_i be the event that ith machine was picked. At time t, let R_t be the event that a machine is running. We know that $P(R_t|M_1) = p_1(t), P(R_t|M_2) = p_2(t), P(M_1) = P(M_2) = 0.5$. Then,

$$P(M_1|R_t) = \frac{P(M_1R_t)}{P(R_t)} = \frac{P(R_t|M_1)P(M_1)}{P(R_t|M_1)P(M_1) + P(R_t|M_2)P(M_2)} = \frac{p_1(t)}{p_1(t) + p_2(t)}$$