1.12 Q4

There are $\binom{20}{10}$ ways of chosing 10 cards from the deck.

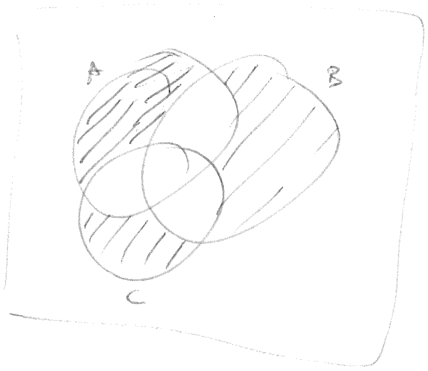
For $j = 1 \ldots 5$ there are 2 ways of chosing two cards with the number $j$.

Hence

$$\frac{\binom{4}{2} \cdots \binom{4}{2}}{\binom{20}{10}} = \frac{6^5}{\binom{20}{10}}.$$

[illegible]

1.12 Q 11




Exactly 1 is the shaded area.

$= P(A) + P(B) + P(C)$

$- 2P(A \cap B) - 2P(A \cap C) - 2P(B \cap C)$

$+ 3P(A \cap B \cap C)$

## 2.5. Q2

a) Sample space

HT   TH

HHT   TTH

HHHT   TTTH

⋮   ⋮

b) exactly 3 tosses.

Pr(HHT or TTH)

mutually exclusive events

$$\Rightarrow \quad \frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} + \frac{1}{2}\times\frac{1}{2}\times\frac{1}{2} = \frac{1}{4}$$

## 2.5 Q5

There are $\binom{10}{3}$ ways of arranging 3 sixes + 7 non-sixes.

One of these has 3 sixes in positions 1, 2, 3

$\Rightarrow$ Prob(1st 3 of 10 rolls are sixes + remainder are non-sixes) $= \dfrac{1}{\binom{10}{3}}$

## 2.5 Q17

$$\Pr(U+V = W+X) = \sum_{j=0}^{18} P(U+V=j)\,P(W+X=j)$$

since U+V is independent of W+X

$P(U+V=j)$ can be found by exhaustively listing all 100 combinations of U, V, each of which has prob. $\frac{1}{100}$

+ counting.

Numerical answer = 0.067.

## 2.5 Q20

$E$ - Event that B wins.

B will win if A misses on 1st try, + B wins on first try

or if both miss on the first try, and B subsequently wins
| ie restart.

$$P(E) = \frac{5}{6} \times \frac{1}{6} + \frac{5}{6} \times \frac{5}{6} \times P(E)$$

$$P(E) = \frac{5}{11}$$

## 2.5 Q21

A wins either if H on 1st toss or A, B, C have T on 1st toss + A subsequently wins

$$P(A) = \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times P(A)$$

$$= \frac{1}{2} + \frac{1}{8} P(A) \qquad P(A) = \frac{4}{7}$$

B wins if A has T + B has H

or All 3 have T + B ~~the~~ subsequently wins

$$P(B) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{8} P(B) \qquad P(B) = \frac{2}{7}$$

$$P(C) = 1 - P(A) - P(B) = \frac{1}{7}.$$

2.5 Q24

Priors $P(A) = 0.2$

$P(B) = 0.5$

$P(C) = 0.3$

$P(\text{lemon} \mid A) = 0.05$

$P(\text{lemon} \mid B) = 0.02$

$P(\text{lemon} \mid C) = 0.1$

$$P(A \mid \text{lemon}) = \frac{P(\text{lemon} \mid A)\,P(A)}{P(\text{lemon})} = \frac{P(\text{lemon} \mid A)\,P(A)}{P(\text{lemon} \mid A)\,P(A) + P(\text{lemon} \mid B)\,P(B) + P(\text{lemon} \mid C)\,P(C)}$$

$$= \frac{0.05 \times 0.2}{0.05 \times 0.2 + 0.02 \times 0.5 + 0.1 \times 0.3} = \frac{1}{5}$$

2.5 Q26

$X$ # tosses on 1st expt

$Y$ # tosses on 2nd expt

$X = n$ if $(n-1)$ T followed by H - occurs with prob $\frac{1}{2^n}$

$Y > n$ if 1st $n$ tosses are T - also occurs with prob $\frac{1}{2^n}$.

$$P(Y > X) = \sum_{n=1}^{\infty} P(Y > n \mid X = n)\,P(X = n)$$

$$= \sum_{n=1}^{\infty} \frac{1}{2^n} \times \frac{1}{2^n} = \sum_{n=1}^{\infty} \frac{1}{4^n} = \frac{1}{3}$$

situation. If $p \neq 1/2$, then it follows from Eq. (2.4.9) that the probabilities $\alpha_1$ and $\alpha_2$ of winning in the two situations equal the values

$$\begin{aligned}
\alpha_1 &= \frac{([1-p]/p)^{50}-1}{([1-p]/p)^{100}-1} = \frac{1}{([1-p]/p)^{50}+1}, \\
\alpha_2 &= \frac{([1-p]/p)^{25}-1}{([1-p]/p)^{50}-1} = \frac{1}{([1-p]/p)^{25}+1}.
\end{aligned}$$

Hence, if $p < 1/2$, then $([1-p]/p) > 1$ and $\alpha_2 > \alpha_1$. If $p > 1/2$, then $([1-p]/p) < 1$ and $\alpha_1 > \alpha_2$.

2.5

36. (a) Since each candidate is equally likely to appear at each point in the sequence, the one who happens to be the best out of the first $i$ has probability $r/i$ of appearing in the first $r$ interviews when $i > r$.

(b) If $i \le r$, then $A$ and $B_i$ are disjoint and $\Pr(A \cap B_i) = 0$ because we cannot hire any of the first $r$ candidates. So $\Pr(A|B_i) = \Pr(A \cap B_i)/\Pr(B_i) = 0$. Next, let $i > r$ and assume that $B_i$ occurs. Let $C_i$ denote the event that we keep interviewing until we see candidate $i$. If $C_i$ also occurs, then we shall rank candidate $i$ higher than any of the ones previously seen and the algorithm tells us to stop and hire candidate $i$. In this case $A$ occurs. This means that $B_i \cap C_i \subset A$. However, if $C_i$ fails, then we shall hire someone before we get to interview candidate $i$ and $A$ will not occur. This means that $B_i \cap C_i^c \cap A = \emptyset$. Since $B_i \cap A = (B_i \cap C_i \cap A) \cup (B_i \cap C_i^c \cap A)$, we have $B_i \cap A = B_i \cap C_i$ and $\Pr(B_i \cap A) = \Pr(B_i \cap C_i)$. So $\Pr(A|B_i) = \Pr(C_i|B_i)$. Conditional on $B_i$, $C_i$ occurs if and only if the best of the first $i-1$ candidates appears in the first $r$ positions. The conditional probability of $C_i$ given $B_i$ is then $r/(i-1)$.

(c) If we use the value $r > 0$ to determine our algorithm, then we can compute

$$p_r = \Pr(A) = \sum_{i=1}^{n} \Pr(B_i)\Pr(A|B_i) = \sum_{i=r+1}^{n} \frac{1}{n}\frac{r}{i-1} = \frac{r}{n}\sum_{i=r+1}^{n}\frac{1}{i-1}.$$

For $r = 0$, if we take $r/r = 1$, then only the first term in the sum produces a nonzero result and $p_0 = 1/n$. This is indeed the probability that the first candidate will be the best one seen so far when the first interview occurs.

(d) Using the formula for $p_r$ with $r > 0$, we have

$$q_r = p_r - p_{r-1} = \frac{1}{n}\left[\sum_{i=r+1}^{n}\frac{1}{i-1} - 1\right],$$

which clearly decreases as $r$ increases because the terms in the sum are the same for all $r$, but there are fewer terms when $r$ is larger. Since all the terms are positive, $q_r$ is strictly decreasing.

(e) Since $p_r = q_r + p_{r-1}$ for $r \ge 1$, we have that $p_r = p_0 + q_1 + \cdots + q_r$. If there exists $r$ such that $q_r \le 0$, then $q_j < 0$ for all $j > r$ and $p_j \le p_{r-1}$ for all $j \ge r$. On the other hand, for each $r$ such that $q_r > 0$, $p_r > p_{r-1}$. Hence, we should choose $r$ to be the last value such that $q_r > 0$.

(f) For $n = 10$, the first few values of $q_r$ are

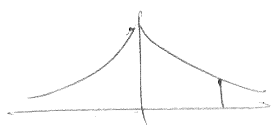
| $r$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $q_r$ | 0.1829 | 0.0829 | 0.0390 | −0.0004 |

So, we should use $r = 3$. We can then compute $p_3 = 0.3987$.

25/10/2016

## 3.11 Q4

$f(x) = \frac{1}{2} e^{-|x|}$



$x_0 > 0$ as $F(x_0) > 0.5$

$\Rightarrow \int_0^{x_0} f(x)\,dx = 0.4$

$\int_0^{x_0} \frac{1}{2} e^{-x}\,dx = 0.4$

$\frac{1}{2} e^{-x} \Big|_0^{x_0} = 0.4$

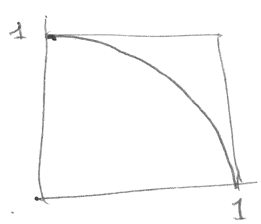
$\frac{1}{2}(e^{-x_0} - 1) = 0.4 \qquad e^{-x_0} = 0.2 \qquad x_0 = \log 5$

## 3.11 Q5

$X_1 \sim U[0,1]$
$X_2 \sim U[0,1]$

$P(X_1^2 + X_2^2 \le 1)$

event is interior of quarter $\bigcirc$.



$P(X_1^2 + X_2^2 \le 1) = \frac{\pi}{4}$

## 3.11 Q8.

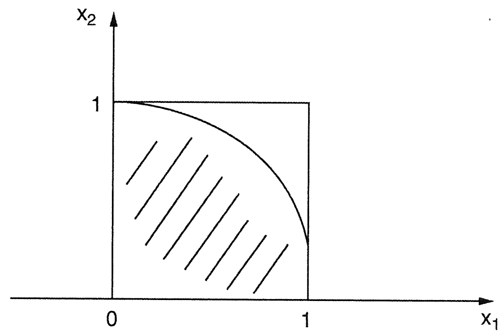



Figure S.3.37: Region for Exercise 5 of Sec. 3.11.

6. (a) $$\Pr(X \text{ divisible by } n) = f(n) + f(2n) + f(3n) + \cdots = \sum_{x=1}^{\infty} \frac{1}{c(p)\ (nx)^p} = \frac{1}{n^p}.$$

(b) By part (a), $\Pr(X \text{ even}) = 1/2^p$. Therefore, $\Pr(X \text{ odd}) = 1 - 1/2^p$.

7.

$$\begin{aligned}\Pr(X + X_2 \text{ even}) &= \Pr(X_1 \text{ even})\Pr(X_2 \text{ even}) + \Pr(X_1 \text{ odd})\Pr(X_2 \text{ odd}) \\ &= \left(\frac{1}{2^p}\right)\left(\frac{1}{2^p}\right) + \left(1 - \frac{1}{2^p}\right)\left(1 - \frac{1}{2^p}\right) \\ &= 1 - \frac{1}{2^{p-1}} + \frac{1}{2^{2p-1}}.\end{aligned}$$

8. Let $G(x)$ devote the c.d.f. of the time until the system fails, let $A$ denote the event that component 1 is still operating at time $x$, and let $B$ denote the event that at least one of the other three components is still operating at time $x$. Then

$$1 - G(x) = \Pr(\text{System still operating at time } x) = \Pr(A \cap B) = \Pr(A)\Pr(B) = [1 - F(x)][1 - F^3(x)].$$

Hence, $G(x) = F(x)\,[1 + F^2(x) - F^3(x)]$.

9. Let $A$ denote the event that the tack will land with its point up on all three tosses. Then $\Pr(A \mid X = x) = x^3$. Hence,

$$\Pr(A) = \int_0^1 x^3 f(x)\,dx = \frac{1}{10}.$$

10. Let $Y$ denote the area of the circle. Then $Y = \pi X^2$, so the inverse transformation is

$$x = (y/\pi)^{1/2} \quad \text{and} \quad \frac{dx}{dy} = \frac{1}{2(\pi y)^{1/2}}.$$

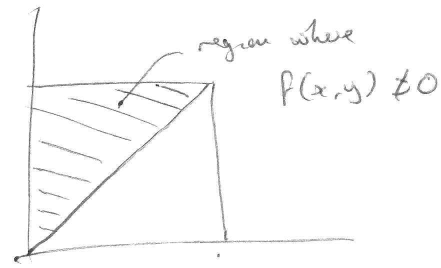
Also, if $0 < x < 2$, then $0 < y < 4\pi$. Thus,

$$g(y) = \frac{1}{16(\pi y)^{1/2}}\left[3\left(\frac{y}{\pi}\right)^{1/2} + 1\right] \quad \text{for} \quad 0 < y < 4\pi$$

and $g(y) = 0$ otherwise.

3.11 Q16.

$$f(x,y) = \begin{cases} 2(x+y) & 0 \leq x < y < 1 \\ 0 & \end{cases}$$




$$f_X(x) = \int_{\cancel{y}\,x}^{1} 2(x+y)\,dy$$

$$= \Big[2xy + y^2\Big]_x^1$$

$$= 2x + 1 - 2x^2 - \cancel{y}x^2$$

$$= 1 + 2x - 3x^2$$

$$P(X < \tfrac{1}{2}) = \int_0^{1/2} (1 + 2x - 3x^2)\,dx = x + x^2 - x^3 \Big|_0^{1/2} = \frac{5}{8}.$$

$$g(y|x) = \frac{f(x,y)}{f(x)} = \frac{2(x+y)}{1+2x-3x^2} \qquad 0 < x < y < 1$$

3.11 Q21

$X \sim e^{-x} \qquad x > 0$

$Y \sim e^{-y} \qquad y > 0$

$$U = \frac{X}{X+Y} \qquad V = X + Y$$

$$f(x,y) = e^{-(x+y)}$$

$x = uv$

$y = (1-u)v \qquad$ Jacobian is $\begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v.$

$$g(u,v) = f(uv, (1-u)v)\,|J| = v\exp(-v) \qquad 0 < u < 1,\ v > 0$$

$g$ can be factored as $g_u(u)g_v(v)$ $\quad (g_u(u) = 1)$ $\Rightarrow$ independent.

3.11 Q27.

$$\begin{array}{cc} & \begin{array}{c}\text{Players in game } n+1\end{array} \\ & \begin{array}{ccc} AB & AC & BC \end{array} \\ \text{Players in game } n. \quad \begin{array}{c} AB \\ AC \\ BC \end{array} & \begin{pmatrix} 0 & 0.3 & 0.7 \\ 0.6 & 0 & 0.4 \\ 0.8 & 0.2 & 0 \end{pmatrix} \end{array}$$

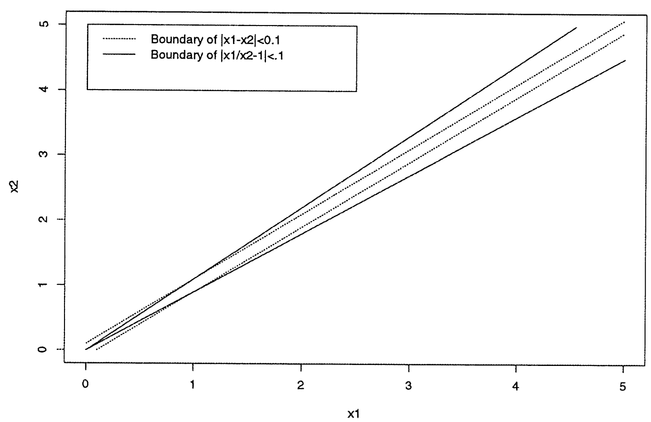



Figure S.3.42: Boundaries of the two regions where $|x_1 - x_2| < 0.1$ and $|x_1/x_2 - 1| < 0.1$ in Exercise 26e of Sec. 3.11.

28. If $A$ and $B$ play in the first game, then there are the following two sequences of outcomes which result in their playing in the fourth game:
i) $A$ beats $B$ in the first game, $C$ beats $A$ in the second game, $B$ beats $C$ in the third game;
ii) $B$ beats $A$ in the first game, $C$ beats $B$ in the second game, $A$ beats $C$ in the third game.

The probability of the first sequence is $(0.3)\ (0.4)\ (0.8) = 0.096$. The probability of the second sequence is $(0.7)\ (0.2)\ (0.6) = 0.084$. Therefore, the overall probability that $A$ and $B$ will play again in the fourth game is 0.18. The same sort of calcuations show that this answer will be the same if $A$ and $C$ play in the first game or if $B$ and $C$ play in the first game.

29. The matrix $\boldsymbol{G}$ and its inverse are

$$\boldsymbol{G} = \begin{pmatrix} -1.0 & 0.3 & 1.0 \\ 0.6 & -1.0 & 1.0 \\ 0.8 & 0.2 & 1.0 \end{pmatrix},$$
$$\boldsymbol{G}^{-1} = \begin{pmatrix} -0.5505 & -0.4587 & 0.5963 \\ 0.0917 & -0.8257 & 0.7339 \\ 0.4220 & 0.2018 & 0.3761 \end{pmatrix}.$$

The bottom row of $\boldsymbol{G}^{-1}$ is the unique stationary distribution, $(0.4220, 0.2018, 0.3761)$.

## 4.9 Q5

$E[X] = \mu.$

$Var(X) = \sigma^2$

$Y = aX + b$

$E[Y] = aE[X] + b = a\mu + b.$

$Var(Y) = a^2 Var(X) = a^2\sigma^2$

$\Rightarrow \quad a\mu + b = 0 \qquad \Rightarrow a = \pm\frac{1}{\sigma}$

$a^2\sigma^2 = 1 \qquad b = -a\mu.$

## 4.9 Q8.

$f(x) = 2x \quad 0 < x < 1$

$Y_n = \max\{X_1 \ldots X_n\} \qquad X_i$ iid $f(x).$

CDF $F_{Y_n}(y) = P(Y_n \leq y) = P($all $X_i$'s are less than $y).$

$$P(X_i < y) = \int_0^y f(x)\,dx = x^2\Big|_0^y = y^2.$$

$$F_{Y_n}(y) = (y^2)^n = y^{2n}$$

$$f_{Y_n}(y) = \frac{d}{dy} y^{2n} = 2ny^{2n-1}$$

$$E[Y_n] = \int_0^1 y_n \times 2ny^{2n-1}\,dy = \frac{2n}{2n+1}$$

4.9 Q11

11. Suppose that you order $s$ liters. If the demand is $x < s$, you will make a profit of $gx$ cents on the $x$ liters sold and suffer a loss of $c(s-x)$ cents on the $s-x$ liters that you do not sell. Therefore, your net profit will be $gx - c(s-x) = (g+c)x - cs$. If the demand is $x \geq s$, then you will sell all $s$ liters and make a profit of $gs$ cents. Hence, your expected net gain is

$$\begin{aligned} E &= \int_0^s [(g+c)x - cs]f(x)dx + gs\int_s^\infty f(x)dx \\ &= \int_0^s (g+c)x\ f(x)dx - csF(s) + gs[1-F(s)]. \end{aligned}$$

To find the value of s that maximizes $E$, we find, after some calculations, that

$$\frac{dE}{ds} = g - (g+c)\ \ F(s).$$

Thus, $\frac{dE}{ds} = 0$ and $E$ is maximized when s is chosen so that $F(s) = g/(g+c)$.

12. Suppose that you return at time $t$. If the machine has failed at time $x \leq t$, then your cost is $c(t-x)$. If the machine has net yet failed ($x > t$), then your cost is $b$. Therefore, your expected cost is

$$E = \int_0^t c(t-x)f(x)dx + b\int_t^\infty f(x)dx = ctF(t) - c\int_0^t xf(x)dx + b[1-F(t)].$$

Hence,

$$\frac{dE}{dt} = cF(t) - bf(t).$$

and $E$ will be maximized at a time $t$ such that $cF(t) = bf(t)$.

13. $E(Z) = 5(3) - 1 + 15 = 29$ in all three parts of this exercise. Also,

$$\mathrm{Var}(Z) = 25\,\mathrm{Var}(X) + \mathrm{Var}(Y) - 10\,\mathrm{Cov}(X,Y) = 109 - 10\,\mathrm{Cov}(X,Y).$$

Hence, $\mathrm{Var}(Z) = 109$ in parts (a) and (b). In part (c),

$$\mathrm{Cov}(X,Y) = \rho\sigma_X\sigma_Y = (.25)(2)(3) = 1.5$$

so $\quad \mathrm{Var}(Z) = 94.$

4.9 Q14

14. In this exercise, $\sum_{j=1}^{n} y_j = x_n - x_0$. Therefore,

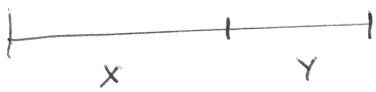
$$\mathrm{Var}(\overline{Y}_n) = \frac{1}{n^2}\mathrm{Var}(X_n - X_0).$$

Since $X_n$ and $X_0$ are independent,

$$\mathrm{Var}(X_n - X_0) = \mathrm{Var}(X_n) + \mathrm{Var}(X_0).$$

Hence, $\mathrm{Var}(\overline{Y}_n) = \dfrac{2\sigma^2}{n^2}$.

4.9 Q22




3 feet long.

break chosen with pdf f(x).

X - length of longer piece

Y - length of shorter piece

$Y = 3 - X$ with prob. 1 $\Rightarrow$ correlation is $-1$

~~$corr(X, 3-X) = -1$~~

5.11 Q7.

$X \sim N(\mu, \sigma^2)$.

$E[X^3]$

consider $E((X-\mu)^3) = E[x^3 - 3\mu x^2 + 3\mu^2 x \mp \mu^3]$

$= E[X^3] - 3\mu E[X^2] + 3\mu^2 E[X] \mp \mu^3$

$E[(X-\mu)^3] = 0$ (odd fn)

$\Rightarrow 0 = E[X^3] - 3\mu E[X^2] + 3\mu^2 E(X] - \mu^3$

$\mathrm{Var}[X] = E[X^2] - (E[X])^2$

$E[X^2] = \mathrm{Var}[X] + (E[X])^2$

$= \sigma^2 + \mu^2$

$\Rightarrow E[X^3] = 3\mu(\sigma^2 + \mu^2) + 3\mu^2\mu - \mu^3$

$= 3\mu\sigma^2 + \mu^3(3 + 3 - 1)$

$= 3\mu\sigma^2 + 5\mu^3$

Book says $-\mu^3$

– where's my error?

5.11 Q8.

$X$ is approx $N(\mu, \frac{12^2}{16})$

$Y$ $N(\mu, \frac{20^2}{25})$

$X - Y$ is approx $N(0, \cancel{16}\ \frac{12^2}{16} + \frac{20^2}{25})$

$N(0, 25)$.

$P(|X-Y| < 5)$

5 is 1 SD $\Rightarrow P(|X-Y|<5) \simeq 0.68$.

5.11 Q12

event $X > n$ : either ~~s~~ $n$ heads or $n$ tails.

$P(X>n) = \left(\frac{1}{2}\right)^n + \left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{n-1}$

this is the CDF.

PMF $P(X=n) = P(X>n-1) - P(X>n)$

$= \left(\frac{1}{2}\right)^{n-2} - \left(\frac{1}{2}\right)^{n-1} = \left(\frac{1}{2}\right)^{n-1}$

$P(X=x) = \left(\frac{1}{2}\right)^{x-1}$ $x = 2, 3, 4 \ldots$

## 5.11 Q20

$X \sim \text{Bin}(n, p)$.

$Y \sim \text{Neg Bin}(r, p)$.

Show that $P(X < r) = \underline{P(Y > n-r)}$

$\uparrow$

less than $r$ successes in $n$ trials.

more than $n-r$ failures before $r^{th}$ success.

$\Rightarrow$ more than $n$ trials are needed to get $r$ successes

## 6.5 Q9.

Exponential dist: $f(x) = \beta e^{-\beta x}$

mean $\frac{1}{\beta}$

variance $\frac{1}{\beta^2}$

mean is $3 \Rightarrow \beta = \frac{1}{3} \Rightarrow$ variance $= 9$.

total time to serve 16 customers.

$\hat{\sim} \; N(16 \times 3, 16 \times 9)$.

$N(48, 144)$.

1 hour $= 60$ mins standardize $\frac{60-48}{\sqrt{144}}$.

$$P(\text{time required} > 60) = 1 - \Phi\left(\frac{60-48}{\sqrt{144}}\right)$$

$$= 1 - \Phi(1)$$

(~~[illegible]~~). $(= 0.1587)$

6.5 Q10.

Poisson $\lambda = 0.01$ per sq. ft.

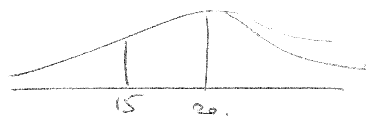
2000 sq ft.

expected # defects $= 20$

Poisson, mean = variance

$\Rightarrow$ variance of # defects $= 20$

# defects $\simeq N(20, 20)$




$P(\text{at least } 15) \simeq 1 - \Phi\left(\frac{15-20}{\sqrt{20}}\right)$

$= 1 - \Phi(-1.180) = 0.8882.$

with continuity correction

$P(\text{at least } 15) \simeq 1 - \Phi\left(\frac{14.5-20}{\sqrt{20}}\right)$

$= 1 - \Phi(1.2298)$

$= 0.8906.$

6.5 Q12

$X \sim NB(n, 0.2)$ n large integer.

a) can use CLT as NB is sum of Geometric ~~Gamma~~ distributed RVs.

$\Rightarrow$ X is sum of large number (n) of iid Geometric ~~Gamma~~ (0.2) RVs.

b) mean

$$E[X] = \frac{r(1-p)}{p} = \frac{n \times 0.8}{0.2} = 4n$$

$$Var(X) = \frac{r(1-p)}{p^2} = \frac{n \times 0.8}{0.2^2} = 20n.$$

$N(4n, 20n)$.