1.

(a)

```
5*|6
6+|34                  n = 20
6*|679                 leaf unit = 0.1 (100kWh)
7+|00234               * for 0-4
7*|566799              + for 5-9
8+|224
```

(b) Negatively skewed.

(c) Min = 5.6 , Max = 8.4

For Q_L, $p=0.25$, $(n+1)p=5.25$,

$$Q_L = X_{(5)} + 0.25(X_{(6)} - X_{(5)}) = 6.7 + 0.25(6.9-6.7) = 6.75$$

For median, $p=0.5$, $(n+1)p=10.5$,

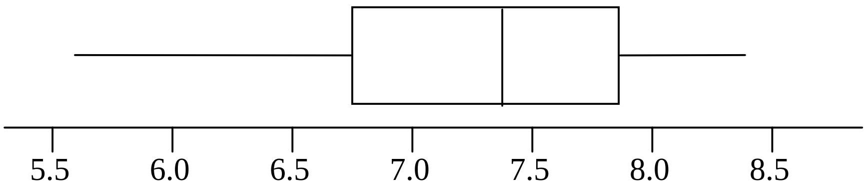
$$meidan = X_{(10)} + 0.5(X_{(11)} - X_{(10)}) = 7.3 + 0.5(7.4-7.3) = 7.35$$

For Q_U, $p=0.75$, $(n+1)p=15.75$,

$$Q_U = X_{(15)} + 0.75(X_{(16)} - X_{(15)}) = 7.7 + 0.75(7.9-7.7) = 7.85$$

Hence the five number summary is : 5.6, 6.75, 7.35, 7.85, 8.4 .

(d) Boxplot for the data :



2. Let X be the power demand in a particular day. Then $X \sim N(15,4)$.

(a) $\Pr(X>17) = 1-\Phi\left(\frac{17-15}{\sqrt{4}}\right) = 1-\Phi(1) = 1-0.8413 = 0.1587$

(b) Let Y be the number of HPD days in six months. Then $Y \sim b(180, 0.1587)$.

$$\Pr(Y \ge 30) \approx 1-\Phi\left(\frac{29.5-(180)(0.1587)}{\sqrt{(180)(0.1587)(0.8413)}}\right) = 1-\Phi(0.19) = 1-0.5753 = 0.4247$$

(c) Let X_i, $i=1,2,\ldots,180$ be the power demand in each of the coming 180 days. Then

$$\overline{X} \sim N\left(15, \frac{4}{180}\right).$$

Total revenue = $0.015\sum_{i=1}^{180} X_i = 2.7\overline{X}$ million dollars.

$$\Pr(\text{Total revenue} > 40 \text{ million dollars}) = \Pr(2.7\overline{X} > 40) = \Pr\left(\overline{X} > \frac{40}{2.7}\right)$$

$$= 1-\Phi\left(\frac{40/2.7-15}{\sqrt{4/180}}\right) = 1-\Phi(-1.242) = \Phi(1.242) = 0.8929$$

2. (d) If the normal assumption is violated, the calculation in (a) will be incorrect. The calculation in (b) relies on the result from (a) and thus is also incorrect. However, the calculation in (c) is still valid because by central limit theorem, the sample mean of a random sample will follow a normal distribution approximately when the sample size is large.

3. $\Pr(I) = 0.3 \qquad \Pr(II) = 0.3 \qquad \Pr(III) = 0.4$

   $\Pr(defective \mid I) = 0.04 \qquad \Pr(defective \mid II) = 0.03 \qquad \Pr(defective \mid III) = 0.02$

   $\Pr(defective) = 0.04 \times 0.3 + 0.03 \times 0.3 + 0.02 \times 0.4 = 0.029$

   (i) $\Pr(I \mid defective) = \dfrac{\Pr(defective \mid I)\Pr(I)}{\Pr(defective)} = \dfrac{0.04 \times 0.3}{0.029} = \dfrac{12}{29} = 0.4138$

   (ii) $\Pr(II \mid defective) = \dfrac{\Pr(defective \mid II)\Pr(II)}{\Pr(defective)} = \dfrac{0.03 \times 0.3}{0.029} = \dfrac{9}{29} = 0.3103$

   (iii) $\Pr(III \mid defective) = \dfrac{\Pr(defective \mid III)\Pr(III)}{\Pr(defective)} = \dfrac{0.02 \times 0.4}{0.029} = \dfrac{8}{29} = 0.2759$

4. Let *X* be the time (in minutes) for a student to finish the examination. Then $X \sim \Gamma(\alpha, \lambda)$.

   (a) $E(X) = 90 \Leftrightarrow \dfrac{\alpha}{\lambda} = 90 \qquad Var(X) = 15^2 \Leftrightarrow \dfrac{\alpha}{\lambda^2} = 225$

   $\lambda = \dfrac{E(X)}{Var(X)} = \dfrac{90}{225} = 0.4 \qquad \alpha = \lambda E(X) = 0.4 \times 90 = 36$

   Hence $X \sim \Gamma(36, 0.4)$.

   (b) Let *c* be the time that should be set. Then $\Pr(X \le c) = 0.9$.

   Since $0.8X \sim \Gamma(36, 0.5) \equiv \chi^2_{72}$,

   $$\Pr(0.8X \le 0.8c) = \Pr(\chi^2_{72} \le 0.8c) = 0.9 \Rightarrow 0.8c \approx 85.53 \Rightarrow c \approx 107 \text{ min}$$

   (Or, $0.8c \approx 85.53 + 0.2(96.58 - 85.53) = 87.74 \Rightarrow c \approx 110$ min)

   (Or, $\Phi\left(\dfrac{0.8c - 72}{\sqrt{144}}\right) \approx 0.9 \Rightarrow \dfrac{0.8c - 72}{12} \approx 1.282 \Rightarrow 0.8c \approx 87.384 \Rightarrow c \approx 110$ min)

5. Joint and marginal distributions of *X* and *Y*:

	Y			
X	**0**	**1**	**2**	**Marginal**
0	0.90	0.03	0.02	0.95
1	0.03	0.01	0.01	0.05
Marginal	0.93	0.04	0.03	1

   (a) $\Pr(X = 0)\Pr(Y = 0) = 0.95 \times 0.93 = 0.8835 \ne 090 = \Pr(X = 0, Y = 0)$

   Hence *X* and *Y* are not independent.

   (b) $E(X) = 0 \times 0.95 + 1 \times 0.05 = 0.05$

   $E(X^2) = 0^2 \times 0.95 + 1^2 \times 0.05 = 0.05 \qquad Var(X) = 0.05 - 0.05^2 = 0.0475$

   $E(Y) = 0 \times 0.93 + 1 \times 0.04 + 2 \times 0.03 = 0.1$

   $E(Y^2) = 0^2 \times 0.93 + 1^2 \times 0.04 + 2^2 \times 0.03 = 0.16 \qquad Var(Y) = 0.16 - 0.1^2 = 0.15$

(c) $E(XY) = (0)(0)(0.90) + (0)(1)(0.03) + \cdots + (1)(2)(0.03) = 0.03$

$Cov(X,Y) = 0.03 - 0.05 \times 0.1 = 0.025$

$$\rho = \frac{0.025}{\sqrt{0.0475 \times 0.15}} = 0.2962$$

(d) Let Z be the cost on replacing mission buttons on a shirt. Then $Z = 0.25X + 0.3Y$.

$E(Z) = 0.25E(X) + 0.3E(Y) = 0.25 \times 0.05 + 0.3 \times 0.1 = 0.0425$

$$\begin{aligned} Var(Z) &= 0.25^2 Var(X) + 0.3^2 Var(Y) + 2(0.25)(0.3)Cov(X,Y) \\ &= 0.25^2(0.0475) + 0.3^2(0.15) + 2(0.25)(0.3)(0.025) = 0.02022 \end{aligned}$$

(e) Let Z_i, $i = 1,2,...,1000$ be the cost on replacing missing buttons in each of the 1000 shirts. By normal approximation,

$$\overline{Z} \overset{\cdot}{\sim} N\left(0.0425, \frac{0.02022}{1000}\right).$$

$$\begin{aligned} \Pr\left(\sum_{i=1}^{1000} Z_i > 40\right) &= \Pr\left(\overline{Z} > 0.04\right) \approx 1 - \Phi\left(\frac{0.04 - 0.0425}{\sqrt{0.02022/1000}}\right) \\ &= 1 - \Phi(-0.556) = \Phi(0.556) = 0.7109 \end{aligned}$$