SAS Probability & Statistics A Solutions

Section 1

To solve questions in this section, you will need to know the properties of the distribution function and density function.

Properties of the cdf

(a) $0 \leq F_X(x) \leq 1$ for all $x \in \mathbb{R}$

(b) $\lim_{x \to \infty} F_X(x) = 1$, $\lim_{x \to -\infty} F_X(x) = 0$

(c) $F_X(x)$ is a non decreasing function.

(d) $F_X(x)$ is right continuous with left limits

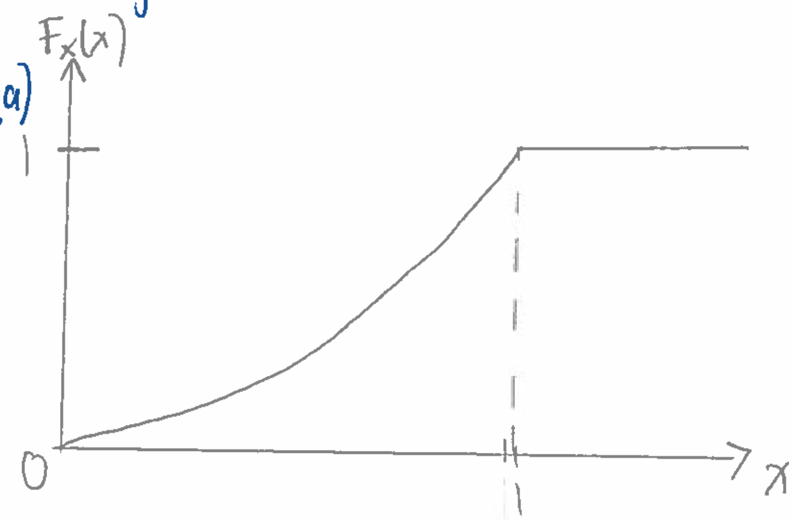
Properties of the density function

(a) The area under the graph for a density function must equal to 1

(b) $f_X(x) \geq 0$ for $x \in \mathbb{R}$

(c) Image of the function is defined at all places where $f_X(x) > 0$

(1)(a)



From the graph above, it is obvious that all the 4 properties for the cdf are statisfied so $F_X(x)$ is a distribution function

To determine the density function:

$$f_X(x) = \begin{cases} \frac{\partial}{\partial x}(x^3) = 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

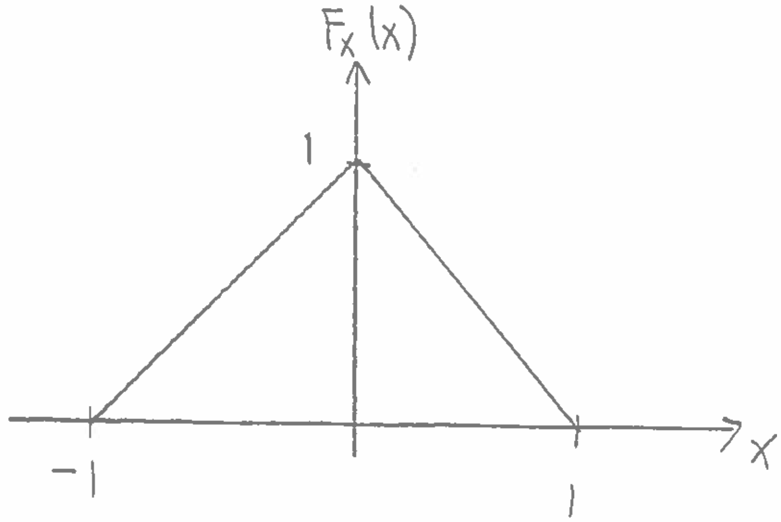
The Image of X is defined when $f_X(x) > 0$

$\text{Im}\, X = [0, 1]$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 x . 3x^2\, dx = \int_0^1 3x^3\, dx$$

$$= \left[\frac{3x^4}{4}\right]_0^1 = \frac{3}{4}$$

(b)

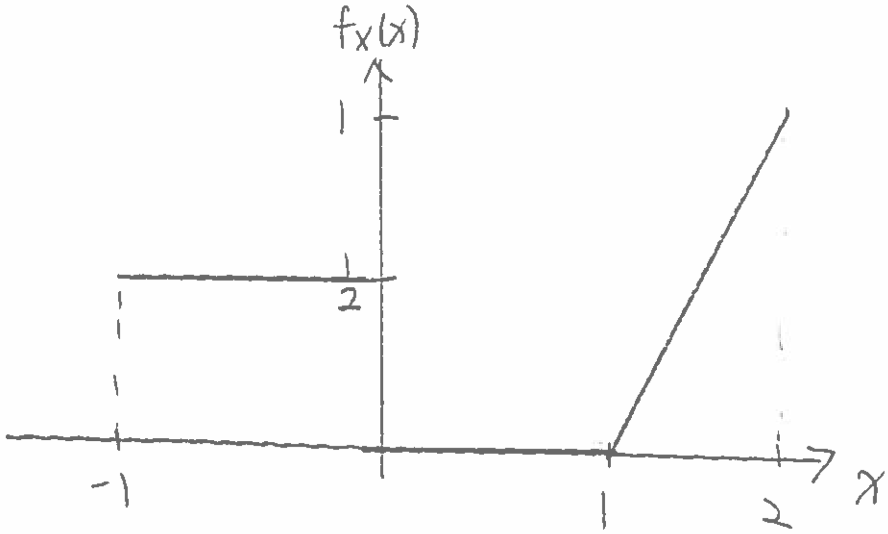


From the graph above, property: $F_X(x)$ is a non decreasing function is not satisfied.

$\therefore$ $F_X(x)$ is not a distribution function

2) The easiest way to check whether a function is a density function is by sketching the graph. To check whether the area under the density function equals 1, it involves integrating the function between the maximum and minimum possible x values it can take.

(a)



Condition (a): Area under the density function $\left[\int_{-\infty}^{\infty} f_X(u)\,du = 1\right]$

This can be shown by a geometric argument or using a calculus argument.

$$\int_{-\infty}^{\infty} f_X(u)\,du = \int_{-1}^{2} f_X(u)\,du = \int_{-1}^{0} \frac{1}{2}\,du + \int_{1}^{2} (u-1)\,du$$

$$= \left[\frac{u}{2}\right]_{-1}^{0} + \left[\frac{u^2}{2} - u\right]_{1}^{2}$$

$$= \frac{1}{2} + \left[\left(\frac{4}{2} - 2\right) - \left(\frac{1}{2} - 1\right)\right]$$

$$= 1$$

Condition (b): $f_X(x) \geq 0$ for $x \in \mathbb{R}$

This can be seen straight away from the graph above

Condition (c): $\operatorname{Im} X = \{x : f_X(u) > 0\}$

$= \{[-1, 0] \cup [1, 2]\}$

Therefore $f_X(x)$ is a density function

To calculate the cdf, the following formula is used:

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(u)\, du$$

The cdf will consist of 5 levels:

$x \in (-\infty, -1]$ $x \in [-1, 0]$, $x \in [0, 1]$, $x \in [1, 2]$ $x \in [2, \infty)$

$\forall x \in (-\infty, -1]$, $F_X(x) = 0$

$$\forall x \in [-1, 0] \Rightarrow F_X(x) = 0 + \int_{-1}^{x} \tfrac{1}{2}\, du = \left[\tfrac{1}{2}u\right]_{-1}^{x} = \tfrac{1}{2}x + \tfrac{1}{2}$$

$$\forall x \in [0, 1] \Rightarrow F_X(x) = \tfrac{1}{2}$$

$$\forall x \in [1, 2] \Rightarrow F_X(x) = \tfrac{1}{2} + \int_{1}^{x} (u-1)\, du = \tfrac{1}{2} + \left[\tfrac{u^2}{2} - u\right]_{1}^{x}$$

$$= 1 + \tfrac{1}{2}x^2 - x$$

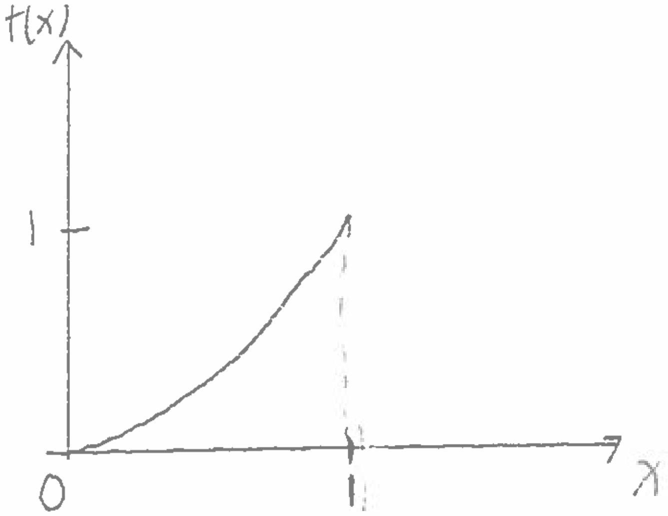
$$\forall x \in [2, \infty) \Rightarrow F_X(x) = 1$$

Putting this all together gives

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \tfrac{1}{2}x + \tfrac{1}{2} & -1 \le x \le 0 \\ \tfrac{1}{2} & 0 \le x \le 1 \\ 1 + \tfrac{1}{2}x^2 - x & 1 \le x \le 2 \\ 1 & x > 2 \end{cases}$$

b) $f_X(x)$ is not a density function because $\int_{-2}^{2} f_X(x)\, dx \neq 1$

c)



Condition (a) $\Rightarrow \displaystyle\int_0^1 3x^2\,dx = \left[x^3\right]_0^1 = 1$ [satisfied]

Condition (b) $\Rightarrow f_X(x) \geq 0$ for $0 < x < 1$ [satisfied]

Condition (c) $\Rightarrow \operatorname{Im} X = \{x : f_X(u) > 0\}$
$= [0, 1]$ [satisfied]

Since all 3 conditions are satisfied, $f(x)$ is a density function

Next is the cdf

$\forall x \in (-\infty, 0]$, $F_X(x) = 0$

$\forall x \in [0, 1]$, $F_X(x) = \displaystyle\int_0^x 3u^2\,du = \left[u^3\right]_0^x = x^3$

$\forall x \in [1, \infty)$, $F_X(x) = 1$

Putting this all together gives

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ x^3 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

Section 2

1) (a) Using $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$,

$$\int_0^1 ax^2\,dx = 1$$

$$\left[\frac{ax^3}{3}\right]_0^1 = 1$$

$$\frac{a}{3} = 1$$

$$a = 3$$

(b) $\forall x \in (-\infty, 0],\ F_X(x) = 0$

$\forall x \in [0,1],\ F_X(x) = \int_0^x 3u^2\,du = \left[u^3\right]_0^x = x^3$

$\forall x \in [1,\infty),\ F_X(x) = 1$

Putting this all together gives

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^3 & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

(c) $E(X) = \int_0^1 x\,3x^2\,dx = \cancel{\left[x^3\right]_0^1 = 1}$

$$= \int_0^1 3x^3\,dx = \left[\frac{3}{4}x^4\right]_0^1 = \frac{3}{4}$$

(d) $Y = X^3$

$$\operatorname{Im} X = [0, 1]$$

$$\operatorname{Im} Y = [0, 1]$$

e) We begin with the distribution function for Y defined as

$$F_Y(y) = P(Y \leq y)$$

You can rewrite Y in terms of X to get the following

$$F_Y(y) = P(Y \leq y) = P(X^3 \leq y) = P(X \leq y^{1/3})$$

$$= F_X(y^{1/3})$$

Substitute $y^{1/3}$ into the cdf of x to get:

$$\Rightarrow (y^{1/3})^3 = y$$

The above is defined for $0 \leq y < 1$

$$\forall y \in (-\infty, 0],\ F_Y(y) = 0$$

$$\forall y \in [1, \infty),\ F_Y(y) = 1$$

Putting this all together gives:

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases}$$

f) $E(Y) = \int_0^1 y \, dy = \left[\frac{y^2}{2}\right]_0^1 = \frac{1}{2}$

Note: If you cannot solve (e), it is still possible to solve (f)

Formula required $E(g(\underline{X})) = \int_{-\infty}^{\infty} g(\underline{X}) f_X(x)\, dx$

Applying this formula gives

$$E(Y) = E(X^3) = \int_0^1 3x^2 . x^3\, dx = \int_0^1 3x^5\, dx$$

$$= \left[\frac{x^6}{2} \right]_0^1 = \frac{1}{2}$$

Both methods will give the same answer.

Section 3

1) (a) It is very important to be aware of what values Y can take when working out the limits or for the integration.

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy$$

$$= \int_0^x 8xy\, dy$$

$$= \left[4xy^2 \right]_{y=0}^{y=x} = 4x^3$$

$$f_X(x) = \begin{cases} 4x^3 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx$

$$= \int_y^1 8xy \; dx$$

$$= \left[4x^2 y\right]_{x=y}^{x=1}$$

$$= 4y - 4y^3 = 4y(1-y^2)$$

$$\therefore f_Y(y) = \begin{cases} 4y(1-y^2) & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(c) Use the following formula for solving conditional density problems:

$$f_{Y|X}(y|x) = \begin{cases} \dfrac{f_{X,Y}(x,y)}{f_X(x)} & f_X(x) > 0 \\ 0 & \text{otherwise} \end{cases}$$

For $f_X(x) > 0$

$$f_{Y|X}(y|x) = \frac{8xy}{4x^3} = \frac{2y}{x^2} \quad \text{for } 0 < y < x < 1$$

(d) To find $P(Y > \frac{1}{4} | X = \frac{3}{4})$,

You will need to determine $f_{Y|X}(y|\frac{3}{4}) = \dfrac{f_{XY}(\frac{3}{4}, y)}{f_X(\frac{3}{4})}$

$$f_{Y|X}(y|\tfrac{3}{4}) = \frac{8 \times \frac{3}{4} \times y}{4 \times (\frac{3}{4})^3} = \frac{6y}{27/16} = \frac{32}{9} y$$

When $x = \frac{3}{4}$, y can only take values $0 \le y \le \frac{3}{4}$

$$P(Y > \tfrac{1}{4} \mid X = \tfrac{3}{4}) = \int_{\frac{1}{4}}^{3/4} f_{Y|X}(y \mid \tfrac{3}{4})\, dy$$

$$= \int_{\frac{1}{4}}^{3/4} \frac{32}{9} y \, dy = \left[\frac{32}{18} y^2\right]_{\frac{1}{4}}^{3/4}$$

$$= \frac{8}{9}$$

(e) $$E(Y \mid X = \tfrac{3}{4}) = \int_0^{\frac{3}{4}} y \, f_{Y|X}(y \mid \tfrac{3}{4})\, dy$$

$$= \int_0^{\frac{3}{4}} y \cdot \frac{32}{9} y \, dy$$

$$= \int_0^{\frac{3}{4}} \frac{32}{9} y^2 \, dy = \left[\frac{32}{27} y^3\right]_0^{3/4} = \frac{1}{2}$$

2) (a) The independence property means $f_{XY}(x,y) = f_X(x) f_Y(y)$

$$f_{XY}(x,y) = 3e^{-3x} e^{-y}$$

(b) First find the distribution function

$$F_Z(z) = P(Z \le z) = P\left(\frac{Y}{X} \le z\right) = P(Y \le zX)$$

Then, do a double integration between $y=0$ and $y = zX$ and $x=0$ and $x=\infty$

$$F_Z(z) = P(Y \le zX) = \int_0^{\infty} \int_0^{y=zX} 3e^{-3x} e^{-y} \, dy \, dx$$

$$= \int_0^{\infty} \left[-3e^{-3x} e^{-y}\right]_0^{zX} dx$$

$$= \int_0^{\infty} -3e^{-3X}(e^{-zX} - 1)\, dx$$

$$= \left[\frac{3e^{-(3+z)X}}{3+z} - e^{-3x}\right]_0^{\infty} = 1 - \frac{3}{3+z}$$

To obtain the density function:

$$f_Z(z) = \frac{\partial}{\partial z}\left(1 - \frac{3}{3+z}\right) = \frac{3}{(3+z)^2}$$

Since $Z = \frac{Y}{X}$, $\operatorname{Im} Z = [0, \infty)$

$$f_Z(z) = \begin{cases} \frac{3}{(3+z)^2} & z > 0 \\ 0 & \text{otherwise} \end{cases}$$

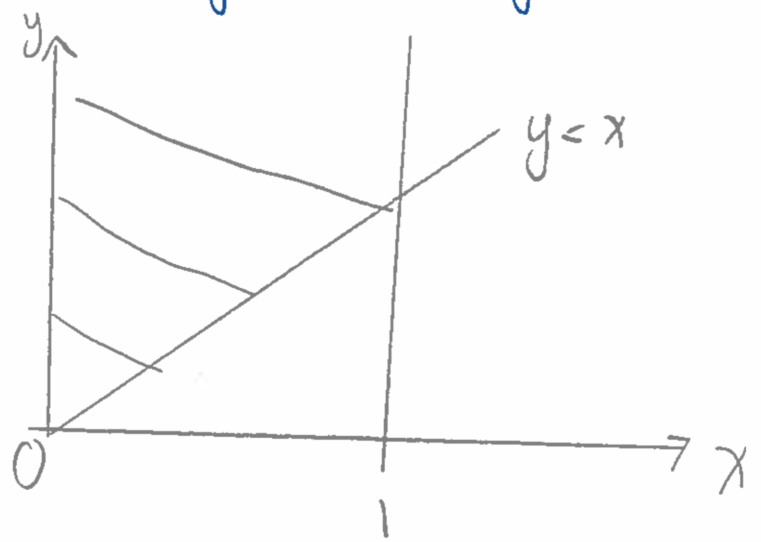
3(a) $f_{XY}(x,y) = f_{Y|X}(y|x)\, f_X(x)$

$$= xe^{-xy} \cdot 2x$$

$$= 2x^2 e^{-xy}$$

$$f_{XY}(x,y) = \begin{cases} 2x^2 e^{-xy} & 0 < x < 1,\ 0 < y < \infty \\ 0 & \text{ow} \end{cases}$$

To calculate $P(Y > X)$, refer to the graph below to determine the shaded region and integral limits



$$P(Y > X) = \int_0^1 \int_x^\infty f_{XY}(x,y)\,dy\,dx$$

$$= \int_0^1 \int_x^\infty 2x^2 e^{-xy}\,dy\,dx$$

$$= \int_0^1 -2xe^{-xy}\Big|_{y=x}^{y=\infty} \cancel{dy}\,dx$$

$$= \int_0^1 2xe^{-x^2}\,dx = 1-e^{-1}$$

(b) $f_{Y|X}(y|x) = \begin{cases} xe^{-xy} & 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$

$f_{Y|X}(y|x)$ has an $Exp(x)$ density

$E(Y|X=x) = \frac{1}{x}$

Section 4

(1) (a) $M_{\bar{X}}(\theta) = E(e^{\bar{X}\theta}) = E\left(e^{\theta\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)}\right)$

~~iid property~~

$\overset{\text{independence}}{=} E\left(e^{\frac{\theta X_1}{n}}\right)E\left(e^{\frac{\theta X_2}{n}}\right)\ldots E\left(e^{\frac{\theta X_n}{n}}\right)$

$\overset{\text{X's are identically distributed}}{=} \left[E\left(e^{\frac{\theta X}{n}}\right)\right]^n = \left(M_X\left(\frac{\theta}{n}\right)\right)^n$

(b) You will need the mgf of a Gamma distribution

mgf of Gamma $\Gamma(t,\lambda) = \left(\frac{\lambda}{\lambda-\theta}\right)^t$

$$M_{\bar{X}}(\theta) = \left(M_X\left(\frac{\theta}{n}\right)\right)^n = \left[\left(\frac{3}{3\cancel{2} - \frac{\theta}{n}}\right)^{2}\right]^n$$

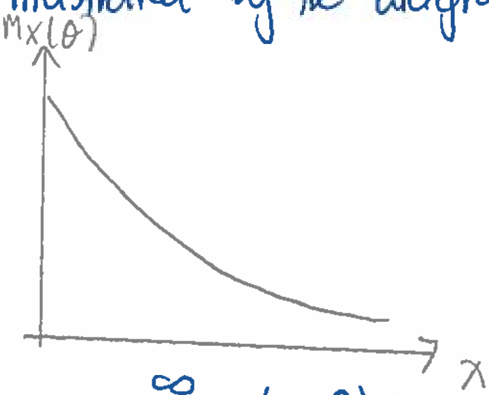
$$= \left(\frac{3}{3\cancel{2} - \frac{\theta}{n}}\right)^{2n} = \left(\frac{3n}{3\cancel{2}n - \theta}\right)^{2n}$$

By the uniqueness of the mgf, $\bar{X} \sim \Gamma(2n, 3n)$

2) (a) $M_X(\theta) = E(e^{\theta X}) = \int_0^\infty e^{\theta x} \lambda e^{-\lambda x}\, dx$ [density of exponential(λ)]

$$= \lambda \int_0^\infty e^{-(\lambda-\theta)x}\, dx$$

For this integral to be finite, we need $\lambda - \theta > 0$. This is illustrated by the diagram below



$$\Rightarrow \lambda \int_0^\infty e^{-(\lambda-\theta)x}\, dx = \lambda \left[\frac{-e^{-(\lambda-\theta)x}}{\lambda-\theta}\right]_0^\infty$$

$$= \frac{\lambda}{\lambda-\theta} \quad \text{if} \quad \theta < \lambda$$

$M_X(\theta)$ is undefined for cases $\lambda - \theta < 0$ and $\lambda - \theta = 0$, because the integral is infinite.

(b) $M_Y(\theta) = E(e^{\theta Y}) = E(e^{c\theta X}) = M_X(c\theta)$

$$M_Y(\theta) = \frac{\lambda}{\lambda - c\theta} = \frac{\lambda/c}{\frac{\lambda}{c} - \theta}$$

This is the mgf of $\text{Exp}\left(\frac{\lambda}{c}\right)$

By the uniqueness of the mgf, $Y \sim \text{Exp}\left(\frac{\lambda}{c}\right)$

Section 5

(1) (a)(i) $X_i \sim \text{Exp}(3)$

By definition

$$M_T(\theta) = E(e^{\theta T}) = E\left(e^{\theta(X_1 + \dots + X_{100})}\right)$$

$$\overset{\text{independence}}{=} E(e^{\theta X_1})\, E(e^{\theta X_2}) \dots E(e^{\theta X_{100}})$$

The X's are identically distributed so,

$$= \left[E(e^{\theta X_1})\right]^n = \left(\frac{\lambda}{\lambda - \theta}\right)^n$$

By the uniqueness of MGF's, $T_{100} \sim \Gamma(n, \lambda)$

$\Rightarrow T_{100} \sim \Gamma(100, 3)$

(ii) Suppose $X_1, \dots, X_i$ are iid RV's with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$ and let $S_n = X_1 + \dots + X_n$

Then for any $a < b$

$$P\left(a \le \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le b\right) \Rightarrow P(a \le Z \le b) \text{ as } n \to \infty$$

where $Z \sim N(0, 1)$

$$P(T_{100} \leq 35) = P\left(\frac{T_{100} - n\mu}{\sqrt{\sigma^2 n}} \leq \frac{35 - n\mu}{\sqrt{\sigma^2 n}}\right)$$

$$n\mu = 100E(X_i) = \frac{100}{3}$$

$$n\sigma^2 = 100 \times Var(X_i) = \frac{100}{9}$$

By the CLT, we get:

$$\approx P\left(Z \leq \frac{35 - \frac{100}{3}}{\sqrt{\frac{100}{9}}}\right) = P(Z \leq \tfrac{1}{2}) \qquad Z \sim N(0,1)$$

$$= 0.6915$$

(b) Note that $Y \sim Bin(150, 0.4)$

To find $P(Y \geq 50)$, apply the CLT

$P(Y \geq 50) = P(Y \geq 49.5) \leftarrow$ Continuity correction is applied here because Y is a discrete distribution

$E(Y) = \cancel{60}\, np = 60 \quad Var(Y) = np(1-p) = 36$

$$\Rightarrow P\left(\frac{Y-60}{6} \geq \frac{49.5-60}{6}\right) \overset{CLT}{\approx} P(Z \geq -\tfrac{7}{4}) = P(Z \leq \tfrac{7}{4})$$

$$\approx 0.95$$

2(a) Repeat the CLT definition mentioned in 1(a)(ii)

Let $T = X_1 + \ldots + X_{50}$ (time to process 50 claims)

$$P(T \leq 120) = P(X_1 + \ldots + X_{50} \leq 120)$$

$$= P\left(\frac{X_1 + \ldots + X_{50} - 125}{\sqrt{62.5}} \leq \frac{120 - 125}{\sqrt{62.5}}\right)$$

$$\overset{CLT}{\approx} P(Z \leq -0.63) = 0.2643$$

$E(T) = 50E(X)$

$= 50 \times \frac{5}{2} = 125$

$Var(T) = 50\,Var(X)$

$= \cancel{62.5}\; 50\left(\frac{5}{4}\right)$

$= 62.5$

(b) $Y \sim Bin(50, 0.3)$

$P(Y > 20) = P(Y \geq 20.5)$ [Use continuity correction as in 1(b)]

$= P\left(\frac{Y-15}{\sqrt{10.5}} \geq \frac{20.5-15}{\sqrt{10.5}}\right)$

$E(Y) = 50(0.3) = 15$ $Var(Y) = 50 \times 0.3 \times 0.7 = 10.5$

$\overset{CLT}{\approx} P(Z \geq 1.70) = 0.0446$