Question (2)

(a) We have $v_i = F^{-1}(u_i)$ for some F^{-1}.

$V = F^{-1}(U)$, V is the random variable generating values $\{v_i\}_{i=1}^{n}$

F^{-1} exists hence F is bijective function — one-to one onto.

Property of Uniform $(0,1)$ U, for $0 \leq p \leq 1$, $\underline{p = P(U \leq p)}$

Hence consider $F(y) = P(U \leq F(y))$, $F(y) \underline{\underline{\in [0,1]}}$.

as $F(y)$ represents probability

Since F^{-1} is bijective,

$$U \leq F(y) \iff F^{-1}(U) \leq F^{-1}(F(y))$$
$$\iff F^{-1}(U) \leq y \iff \underline{\underline{V \leq y}}$$

hence $F(y) = P(V \leq y)$
$$= F_V(y)$$

$\Rightarrow F_V = F$ or distribution of V is same as that of F.

(b) Before Proceedingly, let's note that

$0 \le F \le 1$ (covers all values between them)

as $F(y) = P((Y) \le y)$

and F is an increasing function (can be constant at parts also)

as $F_{(y)} = \int_{-\infty}^{y} f_Y(y)\, dy$ and $\underline{f_Y(y) \ge 0}$.

Now for an increasing function f

$$a \le b \Leftrightarrow f(a) \le f(b) \text{ \& hence } \mathbb{1}(a \le b) = \mathbb{1}(f(a) \le f(b))$$

as $a \le b$ & $f(a) \le f(b)$ will be both true or both false at the same time.

and hence $P(Y \le y) = P(F(Y) \le F(y))$
$= F(y)$

Now $F(y) = z$, $z \in [0,1]$ then

$P(F(Y) \le z) = z \;\; \forall \; z \in [0,1]$.

This is same for a uniform$(0,1)$ r.v.

$P(U \le z) = z \;\; \forall \; z \in [0,1]$

hence $F(Y)$ and the Unif$(0,1)$ have same CDFs (distribution) and are equivalent.

We will use this fact later in steps.

So now we have

$$P(D \geq d) = P\left(\max_x \left| \frac{\sum_{i=1}^{n} 1(Y_i \leq x)}{n} - F(x) \right| \geq d\right)$$

as discussed earlier,

$$1(Y_i \leq x) = 1(F(Y_i) \leq F(x))$$

so

$$= P\left(\max_x \left| \frac{\sum_{i=1}^{n} 1(F(Y_i) \leq F(x))}{n} - F(x) \right| \geq d\right)$$

Substitute $F(x) = y$, as $0 \leq F(x) \leq 1, 0 \leq y \leq 1$

so

$$= P\left(\max_{0 \leq y \leq 1} \left| \frac{\sum_{i=1}^{n} 1(F(Y_i) \leq y)}{n} - y \right| \geq d\right)$$

From my claim earlier, $F(Y_i)$ will be acting as a new uniform random variable $Unif(0,1)$ hence $F(Y_i) = Unif(0,1) = U_i$ (say)

$$= P\left(\max_{0 \leq y \leq 1} \left| \frac{\sum_{i=1}^{n} 1(U_i \leq y)}{n} - y \right| \geq d\right)$$

$$= P(E \geq d)$$

So the previous step is justified only because if I take n-independent uniform variables samples from $[0,1]$ & those from $F(Y_i)$s, they will have same probability for same event (say event being $\geq d$), as $F(Y_i)$s are acting as a new uniform variable, completely randomised now & have same CDFs as a separate uniform random variable and hence the chances of both quantities E & D $\geq d$ is same.

Taking $F(Y_i)$ sample is like taking first lot of n- independent uniform samples and taking U_i sample is like taking another lot of n- independent samples, & prob. for any event for these 2 lots will be same (due to randomness, can't differentiate which is which).

Hence $P(D \geq d) = P(E \geq d)$ proved.

Importance of this equation →

$P(D \geq d)$ is same for whatever distribution F is chosen. It means it is independent of that F. distribution.

This can be useful in checking if the data given for a distribution F indeed is of F or not? If the data belonged to F, then $P(D \geq d)$ must not vary too much from $P(E \geq d)$ computed for a uniform(0,1) variable.

If it varies too much, then we can know that the data does not fit F very well.

Hence it forms a Goodness-to-Fit test.