Variables aléatoires

Exercice 1.

QCM 1 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

- a - La loi d'une variable aléatoire continue affecte à chaque valeur x une valeur entre 0 et $+\infty$
- b - La loi d'une variable aléatoire continue affecte à chaque valeur x une probabilité $\mathrm{P}(x)$
- c - La loi d'une variable aléatoire discrète affecte à chaque valeur x une valeur entre 0 et 1
- d - Une fonction de répartition est une fonction décroissante
- e - Pour toute variable aléatoire continue, la fonction de répartition est une fonction en escalier

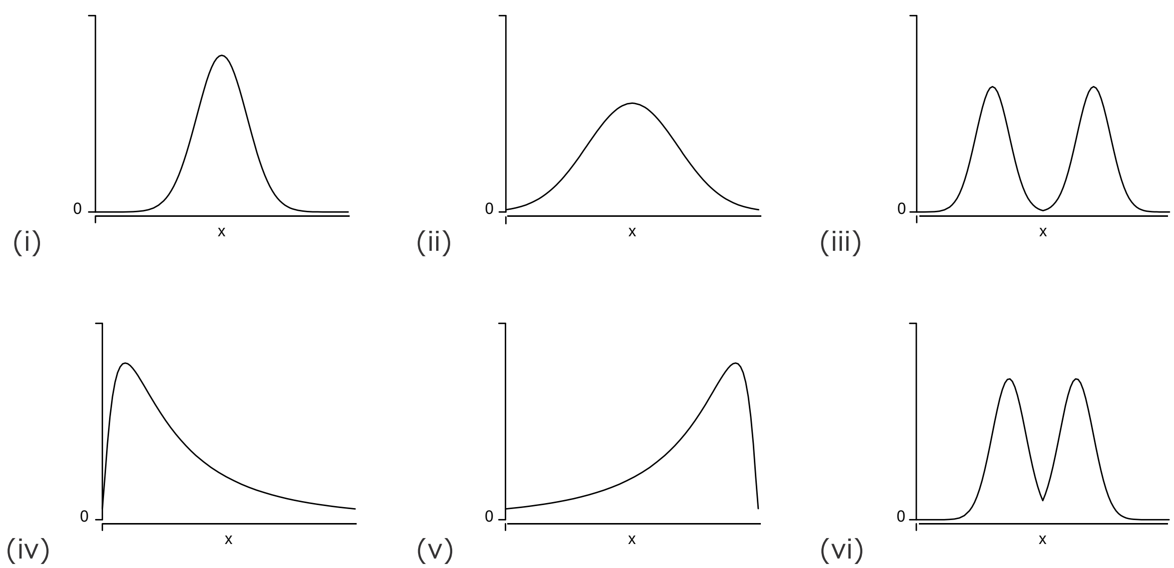
Exercice 2.

QCM 2 Soit X une variable aléatoire continue. Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

- a - La dérivée de la fonction de répartition est la densité de probabilité de cette variable
- b - La probabilité de l'événement $\{X = x\}$ est nulle
- c - Pour les variables aléatoires continues, on ne sait que calculer la probabilité que la réalisation de X soit dans un intervalle autour de x
- d - Une densité de probabilité peut être supérieure à 1
- e - La densité de probabilité affecte à chaque valeur de X une valeur entre 0 et 1

Exercice 3.

Les fonctions suivantes correspondent à six distributions de variable aléatoire (on considère que l'échelle des axes des abscisses est la même pour toutes les courbes).



QCM 3 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

- a - La distribution (i) est symétrique
- b - La distribution (ii) a une moyenne identique à sa médiane
- c - La distribution (iii) a une moyenne inférieure à sa médiane
- d - La distribution (v) a une moyenne supérieure à sa médiane
- e - La distribution (iv) a une moyenne supérieure à sa médiane

QCM 4 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La variable aléatoire correspondant à la distribution (i) a une variance plus élevée que celle de la variable aléatoire de la distribution (ii)

b - La distribution (ii) est plus dispersée que la distribution (i)

c - La variable aléatoire correspondant à la distribution (iii) a une variance plus élevée que celle de la variable aléatoire de la distribution (vi)

d - La distribution (vi) est plus dispersée que la distribution (iii)

e - La fonction (v) est la fonction de répartition de la densité (ii)

Exercice 4.

On considère que, dans chaque famille, les naissances successives sont indépendantes et que la probabilité qu'un garçon naisse est p.

QCM 5 Si $p = 0{,}5$, quelle est la probabilité d'avoir au moins un garçon et une fille dans une famille de 3 enfants ?

a - $< 20\%$

b - Entre 20% et 40%

c - Entre 40% et 60%

d - Entre 60% et 80%

e - $> 80\%$

QCM 6 On considère que $p = 0{,}5$. Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La probabilité qu'il y ait deux garçons exactement dans une famille de quatre enfants est inférieure à $0{,}5$

b - La probabilité qu'il y ait deux garçons exactement dans une famille de quatre enfants est supérieure à $0{,}5$

c - Sachant que l'aîné est un garçon, la probabilité qu'il y ait deux garçons et deux filles dans une famille de quatre enfants est inférieure à $0{,}5$

d - Sachant que l'aîné est un garçon, la probabilité qu'il y ait deux garçons et deux filles dans une famille de quatre enfants est supérieure à $0{,}5$

e - La probabilité qu'il y ait deux garçons et deux filles dans une famille de quatre enfants, sachant que l'aîné est un garçon, est égale à la probabilité qu'il y ait deux garçons et deux filles dans une famille de quatre enfants quel que soit le sexe de l'aîné

Exercice 5.

Une variable aléatoire X suit une loi uniforme continue sur l'intervalle $[0; 10]$.

QCM 7 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La fonction de densité de X vaut $1/10$ pour $x \in [0; 10]$

b - $P(X = 10) = 0{,}1$

c - L'espérance de $Y = 3X + 3$ vaut 18

d - La variance de $Y = 3X + 3$ vaut 28

e - $P(X < 6) = 0{,}6$

QCM 8 On imagine le jeu suivant qui se joue en un ou deux coups par partie. On tire au sort une valeur de X dans une loi uniforme sur $[0; 10]$. Si la première valeur est supérieure ou égale à 6, on gagne 2 euros et la partie s'arrête en un coup. Dans le cas contraire on tire une seconde valeur, si elle est inférieure à 8, on perd 1 euro ; si elle est supérieure ou égale à 8, on gagne 3 euros. Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - L'espérance du nombre de coups par partie est de 1

b - La probabilité de perdre 1 euro par partie est de $0{,}48$

c - L'espérance de gain par partie est de $0{,}68$ euros

d - L'espérance de gain par partie est de $0{,}60$ euros

e - Un sujet décide de jouer à ce jeu jusqu'à ce qu'il ait gagné au moins 5 euros. La probabilité qu'il fasse exactement deux parties est de $0{,}11$ (à 10^{-3} près)

Exercice 6.

Chez un médecin généraliste, en moyenne 5 appels téléphoniques par heure concernent une demande de consultation d'un adulte et en moyenne 9 une demande de consultation d'un enfant.

QCM 9 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - Le nombre d'appels par heure peut être modélisé par une variable aléatoire distribuée selon une loi exponentielle de moyenne 14

b - Le nombre d'appels par heure peut être modélisé par une variable aléatoire distribuée selon une loi de Poisson de moyenne 14

c - Le nombre d'appels par demi heure peut être modélisé par une variable aléatoire distribuée selon une loi exponentielle de moyenne 7

d - Le nombre d'appels par demi heure peut être modélisé par une variable aléatoire distribuée selon une loi de Poisson de moyenne 7

e - Le nombre d'appels par heure est une variable aléatoire discrète

QCM 10 La probabilité pour qu'au plus 4 appels (quel que soit l'âge des malades) surviennent durant une période de 30 minutes est dans l'intervalle :

a - [0; 0,2[

b - [0,2; 0,4[

c - [0,4; 0,6[

d - [0,6; 0,8[

e - [0,8; 1[

Exercice 7.

Le poids d'une certaine population d'étudiants masculins peut être supposé distribué suivant une loi Normale d'espérance 70 kg et d'écart-type 3000 g.

QCM 11 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La probabilité pour que le poids d'un étudiant soit compris entre 67 et 73 kg est de 72% (à 1% près)

b - On peut considérer que, dans cette population, 95% (à 1% près) des individus ont un poids compris entre 64 et 76 kg

c - On peut considérer que, dans cette population, 95% (à 1% près) des individus ont un poids compris entre 64,8 et 77 kg

d - On peut considérer que, dans cette population, 95% (à 1% près) des individus ont un poids inférieur à 75 kg

e - On peut considérer que, dans cette population, 2,5% (à 0,5% près) des individus ont un poids supérieur à 76 kg

Exercice 8.

Soit X_1 une variable aléatoire suivant une loi normale de moyenne $\mu_1 = 2$ et de variance $\sigma_1^2 = 16$. Soit X_2 une variable aléatoire suivant une loi normale de moyenne $\mu_2 = 0$ et de variance $\sigma_2^2 = 9$. On pose $Z_1 = \frac{X_1 - \mu_1}{\sigma_1}$ et $Z_2 = \frac{X_2 - \mu_2}{\sigma_2}$. X_1 et X_2 sont deux variables aléatoires indépendantes

QCM 12 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - Z_1 suit une loi normale centrée réduite

b - Z_1 et Z_2 ne peuvent pas suivre toutes les deux une loi normale centrée réduite

c - $X_1 + X_2$ suit une loi normale de moyenne 2 et d'écart type 5

d - $X_1 - X_2$ suit une loi normale de moyenne 2 et d'écart type 5

e - $X_1 - X_2$ suit une loi normale de moyenne 2 et d'écart type 1

QCM 13 Soit Y = $Z_1^2 + Z_2^2$. Parmi les propositions suivantes, laquelle (lequelles) est (sont) vraie(s)?

- a - On peut appliquer le théorème limite central pour approximer la loi suivie par Y
- b – Y suit une loi normale
- c - Y suit une loi du Chi2 à 2 degrés de liberté
- d - La médiane de Y vaut zéro
- e - La médiane de Y vaut 1,39 à 0,01 près

Exercice 9.

La médiane de survie des patients atteints d'un certain cancer est égale à 12 mois.

QCM 14 Si l'on suppose que la survie des patients atteints de ce cancer suit une loi exponentielle, la moyenne de survie est :

- a - comprise entre 0 et 6 mois
- b - comprise entre 6 mois et 1 an
- c - comprise entre 1 an et 18 mois
- d - comprise entre 18 mois et 2 ans
- e - égale à la médiane de survie

QCM 15 La probabilité qu'un patient, qui est encore vivant à 1 an, décède la 2ème année est :

- a - impossible à déterminer
- b - égale à 0,125
- c - égale à 0,25
- d - égale à 0,5
- e - égale à 0,75

Fluctuations d'échantillonnage, Estimations

Exercice 1.

Dans une population masculine, il existe 50% de fumeurs. On tire au sort un échantillon de n sujets dans cette population.

QCM 1 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

- a - Si $n = 10$, la loi du nombre de fumeurs sur cet échantillon est une loi de Bernoulli de paramètre $\pi = 0{,}50$
- b - Si $n = 10$, la loi du nombre de fumeurs sur cet échantillon est une loi binomiale de paramètres $n = 10$ et $\pi = 0{,}50$
- c - Si $n = 100$, la loi du nombre de fumeurs sur cet échantillon peut être approximée par une loi normale de moyenne $\mu = 50$ et de variance $\sigma = 25$
- d - L'intervalle de fluctuation à 95% de la proportion de fumeurs pour un échantillon de taille $n = 100$ est compris dans l'intervalle $[0{,}4;\ 0{,}6]$
- e - Si l'on voulait avoir une largeur de l'intervalle de fluctuation à 95% de la proportion de fumeurs égale à 0,1, il faudrait diviser par 4 le nombre de sujets dans l'échantillon

QCM 2 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

- a- Si $n=100$, la probabilité que plus de 50 sujets de l'échantillon soient fumeurs est de $0{,}5$
- b- Si $n=10$, la probabilité de trouver moins de 30% de fumeurs dans l'échantillon est de 17% (à 0,5% près)
- c- Si $n=10$, la probabilité de trouver moins de 30% de fumeurs dans l'échantillon est de 5% (à 0,5% près)
- d- Si $n=100$, la probabilité de trouver moins de 30% de fumeurs dans l'échantillon est inférieure à 1%
- e- Si $n=100$, la probabilité de trouver plus de 60% de fumeurs dans l'échantillon est de 2,5% (à 0,5% près)

Exercice 2.

On suppose que, dans une population donnée, le poids des enfants à l'âge de 4 ans est une variable aléatoire distribuée selon une loi Normale de moyenne 16 kg et de variance $3\,kg^2$.

QCM 3 Si l'on prend un enfant au hasard, la probabilité que son poids soit compris entre 13 et 15kg est :

- a - comprise entre 0 et 20%
- b - comprise entre 20% et 40%
- c - comprise entre 40% et 60%
- d - comprise entre 60% et 80%
- e - comprise entre 80% et 100%

QCM 4 On prend au hasard 100 enfants de 4 ans et on note leur poids. Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

- a - M_{100}, la moyenne des poids des 100 enfants, est égale à 16 kg
- b - L'espérance de M_{100} est égale à 16 kg
- c - L'intervalle de fluctuation de M_{100}, avec $\alpha = 0{,}05$, est $[15{,}40;\ 16{,}60]$ (bornes arrondies à 0.05 près)
- d - L'intervalle de fluctuation de M_{100}, avec $\alpha = 0{,}05$, est $[15{,}65;\ 16{,}35]$ (bornes arrondies à 0.05 près)
- e - Il s'agit d'un échantillon représentatif

QCM 5 La taille de l'échantillon d'enfants de 4 ans qu'il faudrait prendre pour que l'intervalle de fluctuation à 95% de la moyenne des poids ait une largeur de 1,4 kg est, à 2 sujets près, de :

- a - 25
- b - 50
- c - 100
- d - 200
- e - 400

Exercice 3.

QCM 6 Un étudiant cherche à estimer la moyenne théorique μ d'une variable aléatoire X à partir d'un n-échantillon $(X_1,...,X_n)$. Il décide :

- Soit de pondérer chaque observation i par i et d'utiliser l'estimateur suivant

$$T_n = \frac{2}{n \times (n+1)} \sum_{i=1}^{n} (i \times X_i)$$

-Soit d'utiliser l'estimateur

$$U_n = \frac{1}{n} \sum_{i=1}^{n} (X_i + \frac{1}{n})$$

On rappelle que $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - Tn est un estimateur biaisé de μ

b - Tn est un estimateur non biaisé de μ

c - Un est un estimateur non biaisé de μ

d -Un est un estimateur asymptomatiquement non biaisé de μ

e-La variance de U_n est plus grande que celle de l'estimateur usuel de μ, $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

Exercice 4.

On s'intéresse à la prévalence de la grippe H1N1 parmi les patients présentant des symptômes grippaux. A partir d'un échantillon de 100 patients, les tests sérologiques retrouvent 20 infections à H1N1. Soit π la proportion théorique d'infection à H1N1.

QCM 7 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - π vaut 20%
b - L'intervalle de confiance à 95% de π vaut [12% - 28%]
c - L'intervalle de confiance à 95% de π vaut [14% - 24%]
d - L'intervalle de confiance à 80% de π vaut [15% - 25%]
e - Les conditions de validité du calcul de l'intervalle de confiance à 95% ne sont pas respectées

QCM 8 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - L'intervalle de confiance de la proportion d'infection à H1N1 est aléatoire
b - L'intervalle de confiance à 99% de π est plus large que l'intervalle de confiance à 95%
c - L'intervalle de confiance à 99% de π est plus étroit que l'intervalle de confiance à 95%
d - Si la proportion de 20% avait été trouvée sur un échantillon de 200 patients, l'intervalle de confiance à 95% de π serait plus étroit que celui calculé à la question 1
e - Si la proportion de 20% avait été trouvée sur un échantillon de 200 patients, l'intervalle de confiance à 95% de π serait plus large que celui calculé à la question 1

Exercice 5.
Au cours d'une étude portant sur 100 sujets diabétiques, on a mesuré la glycémie (taux de glucose dans le sang). Dans cette population, la glycémie peut être considérée comme distribuée selon une loi Normale. Les résultats obtenus sont $\sum x_i = 450$ mmol/l et $\sum x_i^2 = 2049{,}75$ (mmol/l)^2

QCM 9 Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La glycémie moyenne dans la population des sujets diabétiques est 4,5 mmol/l

b - L'estimation sur cet échantillon de la glycémie moyenne dans la population des sujets diabétiques est 4,5 mmol/l

c - L'intervalle de pari à 95% de la glycémie moyenne dans la population des sujets diabétiques est [4,4 ; 4,6] (en mmol/l)

d - L'intervalle de confiance à 95% de la glycémie moyenne dans la population des sujets diabétiques est [4,4 ; 4,6] (en mmol/l)

e - Si la glycémie ne suivait pas une loi Normale, il ne serait pas possible de calculer l'intervalle de confiance à 95% de la glycémie moyenne dans la population des sujets diabétiques

QCM 10 On considère l'intervalle de confiance à 95% de la glycémie moyenne dans la population des sujets diabétiques. Parmi les propositions suivantes, laquelle est vraie ?

a - Il y a 5% de chances que cet intervalle contienne la moyenne de la glycémie dans la population des sujets diabétiques

b - 95% des sujets diabétiques ont une glycémie dans cet intervalle

c - 95% des sujets de l'étude ont une glycémie dans cet intervalle

d - Pour une série d'échantillons tirés dans la population d'étude, il y en a en moyenne 5% pour lesquels l'intervalle de confiance ne contient pas la moyenne de la glycémie dans la population des sujets diabétiques

e - Pour une série d'échantillons tirés dans la population d'étude, il y en a en moyenne 95% pour lesquels l'intervalle de confiance ne contient pas la moyenne de la glycémie dans la population des sujets diabétiques

QCM 11 Quelle devrait être la taille minimale de l'échantillon si on voulait avoir une largeur de l'intervalle de confiance à 95% inférieure à 0,1 mmol/l ?

a - Au moins 100 sujets

b - Au moins 200 sujets

c - Au moins 400 sujets

d - Au moins 800 sujets

e - Au moins 1600 sujets

Exercice 6.
Un test de laboratoire est utilisé pour le diagnostic d'une maladie M d'origine bactérienne. Il consiste à ensemencer une boîte de Petri avec un prélèvement, à mettre en culture l'ensemble, puis à compter le nombre de colonies obtenues à 48h. On considère que le nombre de colonies obtenues suit, chez les sujets sains, une loi de Poisson d'espérance 5 et chez les sujets malades, une loi de Poisson de variance 15.
On désigne par X le nombre de colonies observé dans une boîte de Petri.

QCM 12 Chez un sujet sain, parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La probabilité d'observer exactement 3 colonies dans un boîte de Petri est nulle

b - $P(X \geq 10)$ est inférieure à 5%

c - $P(X \geq 10)$ est comprise entre 5% et 10%

d - La médiane de X est inférieure ou égale à 5

e - La médiane de X est égale à 5 car la loi est symétrique

QCM 13 On fixe un seuil à 9 (c'est à dire que l'on considère que le test est positif si $X > 9$). On rappelle que la sensibilité du test est la probabilité qu'il soit positif chez un sujet malade et sa spécificité la probabilité qu'il soit négatif chez un sujet sain. Parmi les propositions suivantes, laquelle est vraie ?

a - La sensibilité est comprise entre 90% et 95% et la spécificité est supérieure à 95%
b - La sensibilité et la spécificité sont comprises entre 90% et 95%
c - La sensibilité et la spécificité sont supérieures à 95%
d - La sensibilité est supérieure à 95% et la spécificité est comprise entre 90% et 95%
e - On ne peut pas calculer la sensibilité et la spécificité car on ne connait pas la prévalence de la maladie

QCM 14 On précise que la prévalence de la maladie M dans la population d'intérêt est 0,1. On rappelle que la valeur prédictive positive (VPP) de ce test est la probabilité qu'un sujet soit malade si le test est positif et la valeur prédictive négative (VPN), la probabilité qu'un sujet soit sain si le test est négatif. Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - La VPP est supérieure à 75%
b - Si l'on observe $X > 9$ dans une boîte de Petri, la probabilité que le sujet soit malade est supérieure à 75%
c - Si l'on observe $X > 9$ dans une boîte de Petri, la probabilité que le sujet ne soit pas malade est inférieure à 5%
d - La probabilité que le sujet ne soit pas malade, si l'on observe $X > 9$ dans une boîte de Petri, est égale à $(1 - VPN)$
e - Si l'on abaisse le seuil à 8, la sensibilité augmente et la spécificité diminue

QCM 15 On pense que dans cette même population la probabilité d'avoir un test positif est 0,1 (avec un seuil à 9). Parmi les propositions suivantes, laquelle (lesquelles) est (sont) vraie(s) ?

a - C'est nécessairement la vraie valeur
b - On n'a aucun moyen de le vérifier
c - Le calcul montre que c'est une bonne approximation à 3% près
d - La loi exacte du nombre de tests positifs sur un échantillon de taille n est une loi de Poisson, puisque le nombre de colonies suit une loi de Poisson
e - La loi exacte du nombre de tests positifs sur un échantillon de taille n est une loi binomiale

On considère maintenant que la probabilité dans cette population d'avoir un test positif est 0,1.

QCM 16 Soit un échantillon de taille 100 obtenu par tirage au sort. Parmi les propositions suivantes, laquelle est vraie ?

a - L'intervalle de pari à 95% du nombre de tests positifs est compris dans l'intervalle $[6; 14]$
b - L'intervalle de pari à 95% du nombre de tests positifs est compris dans l'intervalle $[4; 16]$
c - L'intervalle de pari à 95% du nombre de tests positifs est compris dans l'intervalle $[9; 11]$
d - On ne peut pas calculer l'intervalle de pari à 95% car les conditions de validité ne sont pas vérifiées aux bornes
e - La probabilité d'observer 10 tests positifs exactement sur l'échantillon est nulle, car la loi suivie par le nombre de positifs est approximativement une loi normale

On constistue maintenant un échantillon représentatif de 64 malades. Pour chaque malade, on note la concentration X d'un antibiotique ayant permis l'éradication de toutes les colonies de la boîte. On obtient $x_i = 640\ UI/l$ et $x_i^2 = 10432\ (UI/l)^2$

QCM 17 Parmi les propositions suivantes, laquelle est vraie ?

a - L'intervalle de pari à 95% de la concentration moyenne nécessaire à l'éradication des bactéries dans la population est [8 ; 12] (en UI/l)
b - L'intervalle de confiance à 95% de la concentration moyenne nécessaire à l'éradication des bactéries dans la population est [8 ; 12] (en UI/l)
c - On peut approximer la loi suivie par la concentration moyenne nécessaire à l'éradication des bactéries dans la population par une loi normale
d - On ne peut pas calculer l'intervalle de confiance car $\sqrt{64} < 30$
e - On ne peut pas calculer l'intervalle de pari car $\sqrt{64} < 30$