Η Κανονική Κατανομή

Μιχάλης Ζαζάνης
Τμήμα Στατιστικής
Οικονομικό Πανεπιστήμιο Αθηνών

1 Η κανονική κατανομή

Μια πραγματική τυχαία μεταβλητή Z ονομάζεται τυποποιημένη κανονική αν η συνάρτηση κατανομής της, $\Phi(x) := P(Z \leq x)$, δίδεται από την σχέση

$$\Phi(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy, \qquad x \in \mathbb{R} \tag{1}$$

με αντίστοιχη πυκνότητα πιθανότητας

$$\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \qquad x \in \mathbb{R}. \tag{2}$$

Αν Z είναι τυποποιημένη κανονική τυχαία μεταβλητή, $\mu \in \mathbb{R}$ και $\sigma > 0$, τότε η τυχαία μεταβλητή

$$X = \mu + \sigma Z$$

είναι κανονική με μέση τιμή $EX = \mu$, διασπορά $Var(X) = \sigma^2$ και πυκνότητα πιθανότητας

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}. \tag{3}$$

Για την κατανομή αυτής της τυχαίας μεταβλητής θα χρησιμοποιούμε τον συμβολισμό $\mathscr{N}(\mu, \sigma^2)$.

2 Ο νόμος των μεγάλων αριθμών και το κεντρικό οριακό θεώρημα

Ο νόμος των μεγάλων αριθμών και το κεντρικό οριακό θεώρημα αποτελούν τους δύο θεμελιώδεις οριακούς νόμους της θεωρίας πιθανοτήτων και ακρογωνιαίους λίθους της στατιστικής επιστήμης. Προκειμένου να δώσουμε επιχειρήματα που δικαιολογούν τον νόμο των μεγάλων αριθμών (και τα οποία με μικρές προσθήκες οδηγούν στην μαθηματική απόδειξή του) ξεκινάμε με την ανισότητα του Chebyshev.

Θεώρημα 1. *Έστω X πραγματική τυχαία μεταβλητή με μέσο $\mu < \infty$ και διασπορά $\sigma^2 < \infty$. Για κάθε $\epsilon > 0$ ισχύει η ανισότητα*

$$P\left(|X-\mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}. \tag{4}$$

Απόδειξη: Θα υποθέσουμε ότι η τυχαία μεταβλητή X είναι συνεχής με πυκνότητα πιθανότητας f. Η απόδειξη για διακριτές τυχαίες μεταβλητές είναι ανάλογη. Ισχύει ότι

$$\begin{aligned}
\sigma^2 &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \\
&= \underbrace{\int_{-\infty}^{\mu-\epsilon} (x-\mu)^2 f(x)dx}_{I_1} + \underbrace{\int_{\mu-\epsilon}^{\mu+\epsilon} (x-\mu)^2 f(x)dx}_{I_2} + \underbrace{\int_{\mu+\epsilon}^{\infty} (x-\mu)^2 f(x)dx}_{I_3}.
\end{aligned}$$

Λαμβάνοντας υπ' όψιν ότι $I_2 \geq 0$ έχουμε

$$\begin{aligned}
\sigma^2 \geq I_1 + I_3 &\geq \int_{-\infty}^{\mu-\epsilon} \epsilon^2 f(x)dx + \int_{\mu+\epsilon}^{\infty} \epsilon^2 f(x)dx \\
&= \epsilon^2 \left(\int_{-\infty}^{\mu-\epsilon} f(x)dx + \int_{\mu+\epsilon}^{\infty} f(x)dx \right) \\
&= \epsilon^2 P\left(|X-\mu| \geq \epsilon\right).
\end{aligned} \tag{5}$$

Στις παραπάνω ανισότητες χρησιμοποιήσαμε το γεγονός ότι για $-\infty < x \leq \mu - \epsilon$ και $\mu + \epsilon \leq x < \infty$ ισχύει η ανισότητα $(x-\mu)^2 \geq \epsilon^2$. Από την (5) προκύπτει η (4). ∎

Στην συνέχεια θα διατυπώσουμε και θα αποδείξουμε τον νόμο των μεγάλων αριθμών. Η διατύπωση που θα δώσουμε είναι η απλούστερη δυνατή και βασίζεται σε μια έννοια σύγκλισης ακολουθιών τυχαίων μεταβλητών που ονομάζεται σύγκλιση κατά πιθανότητα. Η αντίστοιχη μορφή του νόμου των μεγάλων αριθμών ονομάζεται «ασθενής νόμος των μεγάλων αριθμών».

Θεώρημα 2. *Έστω $\{X_i\}$, $i = 1, 2, \ldots$ μια ακολουθία από ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με μέσο $\mu < \infty$ και διασπορά $\sigma^2 < \infty$. Αν $S_n := X_1 + X_2 + \cdots + X_n$ τότε*

$$\frac{1}{n} S_n \xrightarrow{P} \mu \qquad \text{όταν } n \to \infty. \tag{6}$$

Η σύγκλιση κατά πιθανότητα που υποδηλώνεται στην παραπάνω σχέση σημαίνει ότι για κάθε $\epsilon > 0$

$$\lim_{n\to\infty} P\left(\left| \frac{S_n}{n} - \mu \right| \geq \epsilon \right) = 0. \tag{7}$$

Απόδειξη: Μπορούμε να βασιστούμε στην ανισότητα του Chebyshev. Ισχύει ότι

$$E\frac{S_n}{n} = \frac{1}{n}E[X_1 + \cdots + X_n] = \frac{1}{n}(EX_1 + \cdots + EX_n) = \mu.$$

Επίσης, αφού τα X_i είναι ανεξάρτητες τυχαίες μεταβλητές, $\mathsf{Var}(S_n) = \mathsf{Var}(X_1 + \cdots + X_n) = \mathsf{Var}(X_1) + \cdots + \mathsf{Var}(X_n) = n\sigma^2$. Επομένως

$$\mathsf{Var}(\frac{1}{n}S_n) = \frac{\sigma^2}{n}$$

και

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\mathsf{Var}(\frac{1}{n}S_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

Το δεξί μέλος της παραπάνω εξίσωσης, για ϵ σταθερό, καθώς το n τείνει στο άπειρο, τείνει στο μηδέν. ∎

Μια ειδική περίπτωση του παραπάνω νόμου είναι εκείνη κατά την οποία οι τυχαίες μεταβλητές παίρνουν μόνο δύο τιμές, την τιμή 1 (επιτυχία) με πιθανότητα p και την τιμή 0 (αποτυχία) με πιθανότητα $q = 1 - p$. Τότε $\mu = EX_1 = p \times 1 + q \times 0 = p$ και ο νόμος των μεγάλων αριθμών λέει ότι το ποσοστό των επιτυχιών σε n ανεξάρτητες δοκιμές, $\frac{1}{n}S_n$, τείνει στην πιθανότητα επιτυχίας p όταν $n \to \infty$.

Ο νόμος των μεγάλων αριθμών που μόλις είδαμε μας εξασφαλίζει ότι $\frac{S_n}{n} - \mu \to 0$ όταν $n \to \infty$. Θα μπορούσαμε να αναρωτηθούμε πόσο γρήγορα τείνει στο μηδέν αυτή η διαφορά. Στο σχήμα 1 φαίνεται μια πραγματοποίηση της ακολουθίας $\frac{1}{n}S_n$ όταν τα X_i είναι ανεξάρτητες τυχαίες μεταβλητές που παίρνουν τις τιμές 0 και 1 με πιθανότητα 1/2. Περιμένουμε ότι ο λόγος S_n/n θα πλησιάζει το 0,5. Αυτό πράγματι συμβαίνει και το ζητούμενο είναι να προσδιορίσουμε την ταχύτητα με την οποία οι στατιστικές διακυμάνσεις γύρω από τον μέσο πηγαίνουν στο μηδέν. Το κεντρικό οριακό θεώρημα μας λέει, ότι η τάξη μεγέθους του «σφάλματος» $\frac{S_n}{n} - \mu$ είναι $n^{-1/2}$. Επιπλέον, (και αυτό είναι εξαιρετικά ενδιαφέρον!) η κατανομή του σφάλματος οριακά είναι ανεξάρτητη από την αρχική κατανομή των τυχαίων μεταβλητών X_i και είναι κανονική.

Θεώρημα 3 (Κεντρικό Οριακό Θεώρημα). *Έστω X_i, $i = 1, 2, \ldots$ ανεξάρτητες, ισόνομες τυχαίες μεταβλητές με πεπερασμένο μέσο μ και πεπερασμένη διασπορά σ^2. Αν $S_n = X_1 + \cdots + X_n$ τότε*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} Z \qquad \text{όταν } n \to \infty \tag{8}$$

όπου Z τυποποιημένη κανονική τυχαία μεταβλητή. Η έννοια της σύγκλισης «$\xrightarrow{d}$» που ονομάζεται σύγκλιση κατά κατανομή είναι ότι,

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \to P(Z \leq x) = \Phi(x) \qquad \text{για κάθε } x \in \mathbb{R}, \text{ όταν } n \to \infty. \tag{9}$$

Για παράδειγμα έστω ότι οι X_i είναι ομοιόμορφα κατανεμημένες τυχαίες μεταβλητές στο διάστημα $[0,6]$. Η μέση τιμή είναι $\mu = \int_0^6 \frac{1}{6} x dx = 3$, η δεύτερη ροπή είναι $EX_1^2 = \int_0^6 \frac{1}{6} x^2 dx = 12$, και η διασπορά είναι $\sigma^2 = EX_1^2 - \mu^2 = 3$. Συνεπώς το κεντρικό οριακό θεώρημα μας λέει ότι

$$\frac{S_n - 3n}{\sqrt{3n}} \approx \mathcal{N}(0,1)$$

ή ισοδύναμα $S_n \approx 3n + Z\sqrt{3n}$ ή ακόμα $\frac{S_n}{n} - 3 \approx Z\sqrt{\frac{3}{n}}$.

3 Ροπογεννήτριες

Έστω X τυχαία μεταβλητή με συνάρτηση κατανομής $F(x) = P(X \leq x)$ και πυκνότητα πιθανότητας $f(x) = F'(x)$. Η ροπογεννήτρια της X ορίζεται ως

$$M(t) = Ee^{tx} = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

για κάθε t για το οποίο η $M(t)$ ορίζεται δηλαδή το ολοκλήρωμα συγκλίνει.

Παραδείγματα:

i) Έστω ότι η X είναι ομοιόμορφη στο $[0,1]$ με πυκνότητα πιθανότητας $f(x) = 1$ όταν $x \in [0,1]$ και $f(x) = 0$ διαφορετικά. Τότε $M(t) = \int_0^1 e^{tx} dx = (e^t - 1)/t$ για κάθε $t \in \mathbb{R}$.

ii) Έστω ότι η X είναι εκθετικά κατανεμημένη με πυκνότητα πιθανότητας $f(x) = \lambda e^{-\lambda x}$ όταν $x \geq 0$ και 0 όταν $x < 0$. Τότε $M(t) = \int_0^\infty \lambda e^{-x(\lambda - t)} dx = \lambda/(\lambda - t)$ για $t < \lambda$. Όταν $t \geq \lambda$ η ροπογεννήτρια δεν ορίζεται.

iii) Έστω ότι η Z είναι τυποποιημένη κανονική τυχαία μεταβλητή. Τότε η ροπογεννήτριά της υπολογίζεται ως

$$M(t) \;=\; \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tx - x^2/2} dx \;=\; e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(x-t)^2} dx \;=\; e^{\frac{1}{2}t^2}$$

για κάθε $t \in \mathbb{R}$.

iv) Έστω ότι η τ.μ. X είναι κανονική $\mathcal{N}(\mu, \sigma^2)$. Τότε μπορούμε να γράψουμε $X = \mu + \sigma Z$ όπου η Z είναι τυποποιημένη κανονική και επομένως η ροπογεννήτρια συνάρτηση της X είναι

$$Ee^{t(\mu + \sigma Z)} = Ee^{t\mu + t\sigma Z} = e^{t\mu} Ee^{(t\sigma)Z} = e^{t\mu + \frac{1}{2}t^2\sigma^2}. \tag{10}$$

Από τον ορισμό είναι σαφές ότι η ροπογεννήτρια προσδιορίζεται μονοσήμαντα από την αντίστοιχη κατανομή πιθανότητας. Το ακόλουθο θεώρημα, το οποίο παραθέτουμε χωρίς απόδειξη, βεβαιώνει ότι, κάτω από κάποιες τεχνικές προϋποθέσεις, ισχύει και το αντίστροφο.

Θεώρημα 4. *Όταν η ροπογεννήτρια έχει μη τετριμμένο πεδίο ορισμού (δηλ. το πεδίο ορισμού της περιέχει ένα ανοικτό διάστημα) τότε προσδιορίζει μονοσήμαντα την κατανομή πιθανότητας.*

Η ροπογεννήτρια συνάρτηση οφείλει το όνομά της στην ακόλουθη ιδιότητα. Για κάθε $n \in \mathbb{N}$ για το οποίο η ροπή τάξης n μιας τυχαίας μεταβλητής X υπάρχει, δηλαδή $E|X|^n < \infty$, ισχύει ότι

$$EX^n = M^{(n)}(0) \tag{11}$$

δηλαδή η n–οστή ροπή ισούται με την n–οστή παράγωγο της ροπογεννήτριας υπολογισμένη στο 0. Για παράδειγμα, στην περίπτωση της εκθετικής κατανομής, $M(t) = \frac{\lambda}{\lambda - t}$ όπως είδαμε και επομένως $M^{(n)}(t) = n!\frac{\lambda}{(\lambda-t)^{n+1}}$. Συνεπώς οι ροπές της εκθετικής κατανομής δίδονται από την σχέση

$$M^{(n)}(0) = \frac{n!}{\lambda^n}.$$

Δεδομένου ότι $M(t) = E[e^{tX}]$ και ότι $e^{tX} = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!}$ μπορούμε (όταν η ροπογεννήτρια ορίζεται για κάθε t σε κάποια ανοικτή περιοχή του μηδενός) να γράψουμε

$$M(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E[X^n]. \tag{12}$$

Συνεπώς, σε μερικές περιπτώσεις ο απλούστερος τρόπος υπολογισμού των ροπών μιας κατανομής είναι από το ανάπτυγμα Taylor της ροπογεννήτριας στο 0. Έτσι, στην περίπτωση της ομοιόμορφης κατανομής στο μηδέν όπως είδαμε έχουμε ότι

$$M(t) = \frac{e^t - 1}{t} = \sum_{n=1}^{\infty} \frac{t^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{(n+1)!}$$

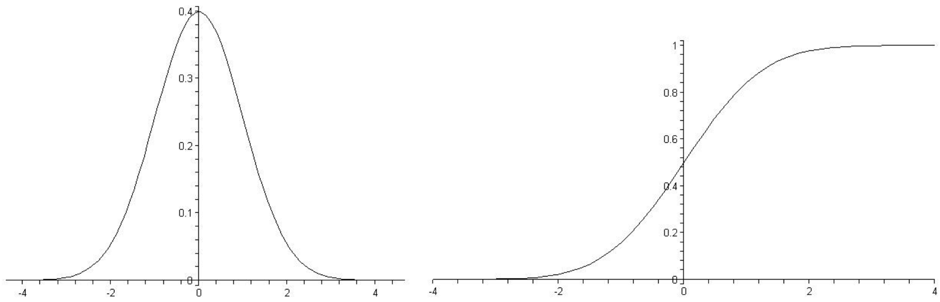
απ' όπου, συγκρίνοντας με την (12) παίρνουμε ότι $EX^n = \frac{1}{n+1}$.

Η περίπτωση της κανονικής κατανομής παρουσιάζει ιδιαίτερο ενδιαφέρον. Έχουμε ότι

$$M(t) = e^{t^2/2} = \sum_{n=0}^{\infty} \frac{t^{2n}}{2^n n!} = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} \frac{(2n)!}{2^n n!}.$$

Συνεπώς, για την τυποποιημένη κανονική κατανομή έχουμε, για $n = 0, 1, 2, \ldots$,

$$\begin{aligned} EX^{2n+1} &= 0 \\ EX^{2n} &= \frac{(2n)!}{2^n n!} = 1 \cdot 3 \cdot 5 \cdots (2n-3) \cdot (2n-1). \end{aligned}$$



Σχήμα 1: Πυκνότητα πιθανότητας και συνάρτηστη κατανομής της τυποποιημένης κανονικής τ.μ.

4 Η λογαριθμοκανονική κατανομή

Έστω Z τυποποιημένη κανονική κατανομή. Η τυχαία μεταβλητή $X = e^Z$ παίρνει μόνο θετικές τιμές και έχει κατανομή που δίνεται, για $x \geq 0$ από την

$$F(x) = P(Z \leq x) = P(Z \leq logx).$$

Η αντίστοιχη πυκνότητα πιθανότητας προκύπτει, παραγωγίζοντας την ανωτέρω σχέση, ως

$$f(x) = \frac{d}{dx}\int_{-\infty}^{\log x} \frac{1}{2\pi} e^{-y^2/2} dy = \frac{1}{x\sqrt{2\pi}} e^{-(\log x)^2/2}, \qquad x > 0. \tag{13}$$

Αυτή η πυκνότητα πιθανότητας ονομάζεται λογαριθμοκανονική και η γραφική της παράσταση φαινεται στο ακόλουθο σχήμα. Η γενική λογαριθμοκανονική κατανομή προκύπτει αν

$$X = e^{\mu+\sigma Z}, \tag{14}$$

δηλαδή ως το εκθετικό μιας κανονικής κατανομής με μέσο μ και τυπική απόκλιση $\sigma > 0$. Η αντίστοιχη πυκνότητα πιθανότητας δίνεται από την σχέση

$$f(x) = \frac{x\sigma\sqrt{2\pi}^{-\frac{1}{2}(\frac{\log x-\mu}{\sigma})^2}}{e}, \qquad x > 0. \tag{15}$$

Οι ροπές της λογαριθμοκανονικής κατανομής υπολογίζονται εύκολα από την ροπογεννήτρια της κανονικής κατανομής ως εξής: Αν η X δίδεται από την σχέση (14) τότε

$$EX^n = E[e^{n\mu+n\sigma Z}] = e^{n\mu}E[e^{n\sigma Z}] = e^{n\mu}e^{\frac{1}{2}n^2\sigma^2} = e^{n\mu+\frac{1}{2}n^2\sigma^2}.$$

Συνεπώς

$$\begin{aligned} EX &= e^{\mu+\frac{1}{2}\sigma^2} \\ EX^2 &= e^{2\mu+2\sigma^2} \end{aligned}$$

$$\mathsf{Var}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

Η λογαριθμοκανονική κατανομή εμφανίζεται σε πολλές εφαρμογές όπου μια ποσότητα προκύπτει με πολλαπλασιασμό πολλών ανεξαρτήτων παραγόντων. Για παράδειγμα, ας υποθέσουμε ότι μια μετοχή ξεκινάει από μια τιμή 100 ευρώ και κάθε μέρα η αξία της μεταβάλεται είτε +1% είτε -1% με πιθανότητα 1/2 ανεξάρτητα από τις προηγούμενες. Μας ενδιαφέρει να προσδιορίσουμε την κατανομή της μετοχής μετά από 100 ημέρες.

5 Διμεταβλητές κατανομές

Δύο πραγματικές τυχαίες μεταβλητές, X, Y, ορισμένες στον ίδιο χώρο πιθανότητας, μπορούν να θεωρηθούν ως ένα τυχαίο σημείο στο επίπεδο $\mathbb{R}^2$. Η από κοινού κατανομή των X, Y, ορίζεται μέσω της συνάρτησης κατανομής $F(x,y) := P(X \leq x, Y \leq y)$. Η $F(x,y)$ είναι αύξουσα ως προς κάθε ένα από τα δύο ορίσματα της και ικανοποιεί τις σχέσεις

$$\begin{aligned}
&\lim_{x\to-\infty} F(x,y) = 0 \quad \text{για κάθε } x \in \mathbb{R},\\
&\lim_{y\to-\infty} F(x,y) = 0 \quad \text{για κάθε } y \in \mathbb{R},\\
&\lim_{\substack{x\to\infty\\ y\to\infty}} F(x,y) = 1.
\end{aligned}$$

Επιπλέον, για κάθε $x_1 \leq x_2$, $y_1 \leq y_2$, ικανοποιεί την ανισότητα

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \geq 0.$$

Οι συναρτήσεις $F_X(x) := \lim_{y\to\infty} F(x,y)$, $F_Y(y) := \lim_{x\to\infty} F(x,y)$, μπορεί να δείξει κανείς εύκολα ότι είναι συναρτήσεις κατανομής και ονομάζονται περιθώριες κατανομές των τ.μ. X και Y αντίστοιχα. Ισχύει ότι $F_X(x) = P(X \leq x)$ και $F_Y(y) = P(Y \leq y)$ αντίστοιχα.

Αν η $F(x,y)$ είναι απολύτως συνεχής τότε η μεικτή μερική παράγωγος

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial y \partial x}$$

υπάρχει και ισχύει ότι

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u,v)dudv.$$

Η $f(x,y)$ ονομάζεται διμεταβλητή πυκνότητα πιθανότητας. Ισχύει ότι

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(u,v)dudv.$$

Επίσης οι περιθώριες πυκνότητες πιθανότητας είναι

$$\begin{aligned} F'_X(x) &= f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy, \\ F'_Y(y) &= f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx. \end{aligned}$$

Τέλος η δεσμευμένη πυκνότητα πιθανότητας του Y δεδομένου ότι $X = x$ δίδεται από τη σχέση

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}. \tag{16}$$

Οι τυχαίες μεταβλητές X, Y, είναι ανεξάρτητες αν $F(x,y) = F_X(x)F_Y(y)$ ή, ισοδύναμα, αν $f(x,y) = f_X(x)f_Y(y)$, για κάθε $x, y \in \mathbb{R}$.

Η από κοινού ροπογεννήτρια των τυχαίων μεταβλητών X, Y, ορίζεται ως η

$$M(u,v) := Ee^{uX+vY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{ux+vy} f(x,y)dxdy \tag{17}$$

για όλες τις τιμες $(u,v) \in \mathbb{R}^2$ για τις οποίες το ολοκλήρωμα συγκλίνει. Όπως και σε μία διάσταση ισχύει το ακόλουθο θεώρημα.

Θεώρημα 5. *Σε κάθε διμεταβλητή κατανομή αντιστοιχεί μια μοναδική ροπογεννήτρια συνάρτηση και σε κάθε ροπογεννήτρια συνάρτηση στον $\mathbb{R}^2$ μια μοναδική διμεταβλητή κατανομή υπό την προϋπόθεση ότι η ροπογεννήτρια ορίζεται σε μια ανοικτή περιοχή του* $(0,0)$.

Έστω X, Y, τυχαίες μεταβλητές με από κοινού κατανομή $F(x,y)$ και $Z = aX + bY$ όπου $a, b \in \mathbb{R}$. Η κατανομή της τ.μ. Z δίδεται από την σχέση $F_Z(z) = P(aX + bY \leq z)$ και η αντίστοιχη πυκνότητα πιθανότητας είναι

$$f_Z(z) = \int_{-\infty}^{\infty} f(x,(z-ax)/b)dx$$

Θεώρημα 6 (Cramér–Wold)**.** *Η γνώση της κατανομής της τυχαίας μεταβλητής $Z = aX + bY$ για κάθε $a, b \in \mathbb{R}$ προσδιορίζει πλήρως την από κοινού κατανομή των X, Y.*

Απόδειξη: Έστω $\Phi(u,v) := Ee^{uX+vY}$ η από κοινού ροπογεννήτρια των X, Y. Έστω επίσης $M_{a,b}(t) := Ee^{tZ} = Ee^{t(aX+bY)}$ η ροπογεννήτρια της Z. Εφ' όσον $\Phi(u,v) = M_{u,v}(1)$, η γνώση της κατανομής της Z, και επομένως και της ροπογεννήτριάς της, για κάθε a, b, προσδιορίζει πλήρως την από κοινού ροπογεννήτρια των X, Y. Αρκεί επομένως να επικαλεσθούμε το θεώρημα (5) για να ολοκληρώσουμε την απόδειξη. ∎

6 Η διμεταβλητή κανονική κατανομή

Ορισμός 1. *Έστω (X, Y) ένα τυχαίο στοιχείο του $\mathbb{R}^2$ (τυχαίο διάνυσμα) και $f(x, y)$ η από κοινού πυκνότητα πιθανότητας του. Η f θα ονομάζεται* μη εκφυλισμένη *διμεταβλητή πυκνότητα πιθανότητας αν*

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\ e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2}-2\rho\frac{x-\mu_x}{\sigma_x}\frac{y-\mu_y}{\sigma_y}+\frac{(x-\mu_x)^2}{\sigma_x^2}\right)}. \tag{18}$$

Οι παράμετροι αντιστοιχούν στους μέσους, τις διασπορές και τον συντελεστή συσχέτισης: $EX = \mu_x$, $EY = \mu_y$, $Var(X) = \sigma_x^2$, $Var(Y) = \sigma_y^2$, και $\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$. Στην ανωτέρω σχέση $-1 < \rho < 1$.

Στον παραπάνω ορισμό αποκλείουμε την περίπτωση $\rho = \pm 1$ και γι' αυτό χαρακτηρίζουμε την διμεταβλητή κατανομή ως μη εκφυλισμένη.

Ορισμός 2. *Οι τυχαίες μεταβλητές (X, Y) θα ονομάζονται από κοινού κανονικά κατανεμημένες (ή από κοινού κανονικές) αν είτε*

ι) η από κοινού πυκνότητα πιθανότητάς τους είναι η μη εκφυλισμένη διμεταβλητή κατανομή (18), ή

ιι) υπάρχουν $a, b \in \mathbb{R}$ έτσι ώστε $Y = aX + b$ και η X είναι κανονικά κατανεμημένη.

Παρατήρηση: Στην περίπτωση ιι) του ορισμού, αν $a \neq 0$ και $X \sim \mathscr{N}(\mu, \sigma^2)$, τότε $Y \sim \mathscr{N}(a\mu + b + a^2\sigma^2)$ και $Cov(X, Y) = a\sigma^2$. Συνεπώς ο συντελεστής συσχέτισης είναι

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{a\sigma^2}{\sqrt{a^2\sigma^4}} = \frac{a}{|a|} = \begin{cases} +1 & \text{αν } a > 0 \\ -1 & \text{αν } a < 0 \end{cases}$$

Αν $a = 0$ τότε $P(Y = b) = 1$ και $Var(Y) = 0$, συνεπώς ο συντελεστής συσχέτισης δεν ορίζεται.

Το ακόλουθο θεώρημα δίνει ένα εναλλακτικό χαρακτηρισμό της απο κοινού κανονικής κατανομής:

Θεώρημα 7. *Οι τυχαίες μεταβλητές (X, Y) είναι από κοινού κανονικά κατανεμημένες αν η από κοινού ροπογεννήτριά τους δίδεται από την σχέση*

$$E[e^{uX+vY}] = e^{\frac{1}{2}\left(u^2\sigma_x^2+2uv\rho\sigma_x\sigma_y+v^2\sigma_y^2\right)+u\mu_x+v\mu_y} \tag{19}$$

όπου $\sigma_x, \sigma_y \geq 0$, $-1 \leq \rho \leq 1$ και $\mu_x, \mu_y \in \mathbb{R}$.

Απόδειξη: Το ευθύ, δηλαδή ότι όταν τα X, Y, είναι κανονικά κατανεμημένα τότε η ροπογεννήτρια Ee^{uX+vY} δίδεται από την σχέση (19) προκύπτει από τον ορισμό. Αν η κανονική κατανομή είναι μη εκφυλισμένη τότε

$$Ee^{uX+vY} = \iint_{\mathbb{R}^2} e^{ux+vy} f(x,y)dxdy$$

όπου η $f(x,y)$ δίδεται από την σχέση (18) και η έκφραση (19) προκύπτει από τον υπολογισμό του διπλού ολοκληρώματος. Αν η κανονική κατανομή είναι εκφυλισμένη τότε $Y = aX + b$ για κάποιες τιμές των a, b, και

$$Ee^{uX+vY} = Ee^{uX+avX+bv} = e^{bv}E^{(u+av)X} = e^{bv}e^{(u+av)\mu_x+\frac{1}{2}(u+av)^2\sigma_x^2}.$$

Η τελευταία έκφραση υπολογίζεται από την έκφραση για την ροπογεννητρια της κανονικής τ.μ. X και ισοδυναμεί με την (19) όταν η από κοινού κατανομή των X, Y, είναι εκφυλισμένη κανονική.

Το αντίστροφο, δηλαδή ότι η (19) συνεπάγεται την από κοινού κανονικότητα των X, Y, είναι αποτέλεσμα του θεμελιώδους θεωρήματος σύμφωνα με το οποίο η ροπογεννήτρια προσδιορίζει μονοσήμαντα την κατανομή. ∎

Για την γενική διμεταβλητή κατανομή θα χρησιμοποιούμε τον συμβολισμό

$$\mathscr{N}\left(\begin{pmatrix}\mu_x\\ \mu_y\end{pmatrix}, \begin{pmatrix}\sigma_x^2 & \rho\sigma_x\sigma_y\\ \rho\sigma_x\sigma_y & \sigma_y^2\end{pmatrix}\right). \tag{20}$$

Θεώρημα 8. *Δύο τυχαίες μεταβλητές X, Y είναι από κοινού κανονικές αν και μόνο άν για κάθε $a, b \in \mathbb{R}$ η $aX + bY$ είναι κανονική τ.μ.*

Απόδειξη: Έστω ότι οι τυχαίες μεταβλητές X, Y έχουν την από κοινού κανονική κατανομή (20). Τότε η $W = aX + bY$ έχει ροπογεννήτρια

$$\begin{aligned} M(t) &= Ee^{tW} = E[e^{t(aX+bY)}] = Ee^{(ta)X+(tb)Y} = E[e^{(ta)\mu_x+(tb)\mu_y+\frac{1}{2}((ta)^2\sigma_x^2+2\rho\sigma_x\sigma_y(ta)(tb)+\sigma_y^2(tb)^2)}] \\ &= E[e^{t(a\mu_x+b\mu_y)+\frac{t^2}{2}(a^2\sigma_x^2+2\rho\sigma_x\sigma_y ab+\sigma_y^2b^2)}] \end{aligned}$$

όπου στον παραπάνω υπολογισμό χρησιμοποιήσαμε την (19). Επομένως, από την (10) προκύπτει ότι η W είναι $\mathscr{N}\left(a\mu_x + b\mu_y, a^2\sigma_x^2 + 2\rho\sigma_x\sigma_y ab + \sigma_y^2 b^2\right)$.

Αντίστροφα, έστω ότι η $W = aX + bY$ είναι κανονικά κατανεμημένη για κάθε $a, b \in \mathbb{R}$. Ασχέτως της από κοινού κατανομής των X, Y, $EW = aEX + bEY$ και $Var(W) = a^2Var(X) + 2abCov(X,Y) + b^2Var(Y)$. Συνεπώς,

$$Ee^{tW} = e^{tEW+\frac{1}{2}t^2Var(W)}$$

ή
$$Ee^{taX+tbY} = e^{taEX+tbEY+\frac{1}{2}((ta)^2Var(X)+2(tatb)Cov(X,Y)+(tb)^2Var(Y))}.$$

Θέτωντας $u = ta$, $v = tb$ στην παραπάνω σχέση βλέπουμε ότι οι X, Y, είναι από κοινού κατανονικά κατανεμημένες. ∎

7 Μια χρήσιμη γεωμετρική αναπαράσταση

Έστω U, V κανονικές τυποποιημένες ανεξάρτητες τυχαίες μεταβλητές και $\rho \in [-1, 1]$. Οι μεταβλητές

$$\begin{aligned} X &= U \\ Y &= \rho U + \sqrt{1-\rho^2}V \end{aligned} \tag{21}$$

είναι από κοινού κανονικές με κατανομή

$$\mathscr{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right). \tag{22}$$

Η παραπάνω αναπαράσταση δίνει ένα εύκολο τρόπο να υπολογίσουμε την πιθανότητα $P(X > 0, Y > 0)$ όταν οι X, Y έχουν την διμεταβλητή τυποποιημένη κατανομή (22). Πράγματι, για $|\rho| < 1$ έχουμε

$$\begin{aligned} P(X > 0, Y > 0) &= P\left(U > 0, \rho U + \sqrt{1-\rho^2}V > 0\right) \\ &= P\left(U > 0, V > -\frac{\rho}{\sqrt{1-\rho^2}}U\right) \\ &= \frac{\pi}{4} + \frac{1}{2\pi}\arcsin\rho. \end{aligned} \tag{23}$$

8 Δεσμευμένες κατανομές

Έστω ότι οι τυχαίες μεταβλητές X, Y, είναι από κοινού κανονικές. Τότε η δεσμευμένη κατανομή του Y δεδομένου ότι $X = x$ είναι κανονική με μέσο

$$E[Y|X = x] = \mu_y + \frac{\rho\sigma_y}{\sigma_x}(x - \mu_x) \tag{24}$$

και διασπορά.

$$Var(Y|X = x) = \sigma_y^2(1 - \rho^2) \tag{25}$$

Παρατηρήστε ότι η δεσμευμένη διασπορά δεν εξαρτάται από την τιμή της X ενώ ο δεσμευμένος μέσος εξαρτάται γραμμικά από το x.

Η απόδειξη της προηγούμενης πρότασης είναι απλή. Αρκεί να χρησιμοποιήσουμε τον ορισμό της δεσμευμένης πυκνότητας πιθανότητας

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

όπου $f_X(x) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-(x-\mu_x)^2/2\sigma_x^2}$ είναι η περιθώρια κατανομή της X που αντιστοιχεί στην από κοινού κατανομή (18). Με βάση τα ανωτέρω προκύπτει ότι

$$f_{Y|X}(y|x) = \frac{1}{\sigma_y\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2\sigma_y^2(1-\rho^2)}(y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x))^2}.$$

9 Προβλήματα

1. Έστω X, Y, δύο από κοινού κανονικές τυχαίες μεταβλητές.

   α) Να δείξετε ότι οι τυχαίες μεταβλητές $X+Y$, $X-Y$, είναι ασυσχέτιστες (και επομένως και ανεξάρτητες) αν και μόνο αν $Var(X) = Var(Y)$.

   β) Αν επιπλέον $Var(X) = Var(Y) = 1$ και $EX = EY = 0$, και ο συντελεστής συσχέτισης μεταξύ των X και U είναι $\rho = 0.6$, να υπολογίσετε την πιθανότητα $P(X-Y<1, X+Y>2)$.

2. Έστω (X,Y) από κοινού κανονικές τυχαίες μεταβλητές με κατανομή

$$\mathcal{N}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}\right).$$

   α) Να υπολογίσετε την πυκνότητα πιθανότητας της τυχαίας μεταβλητής $W = X+2Y$.

   β) Να υπολογίσετε την δεσμευμένη μέση τιμή $E[Y|X=3]$.

   γ) Να υπολογίσετε την μέση τιμή $E[W^2]$.

   δ) Αν $V = 2X - Y$, ποια είναι η από κοινού κατανομή των W, V;

   ε) Ποια είναι η δεσμευμένη μέση τιμή $E[W|V=0]$;

3. Έστω X, Y, δύο ανεξάρτητες, τυποποιημένες κανονικές τυχαίες μεταβλητές.

α) Να υπολογίσετε την πιθανότητα $P(X \geq 0, Y \geq 0)$.

β) Να υπολογίσετε την πιθανότητα $P(|X+Y| \leq 1)$.

γ) Να υπολογίσετε την πιθανότητα $P(|X+Y| \leq 1, |X-Y| \leq 1)$.

δ) Να υπολογίσετε την πιθανότητα $P(X \geq 0, Y \geq 0, |X+Y| \leq 1)$.

4. Έστω X, Y, δύο ανεξάρτητες, τυποποιημένες κανονικές τυχαίες μεταβλητές.

α) Να υπολογίσετε την πιθανότητα $P(X > 2Y)$.

β) Αν $U = X\sqrt{3} + Y$, $V = X - Y\sqrt{3}$ να υπολογίσετε την πιθανότητα $P(U > 2V)$.

γ) Να υπολογίσετε την πιθανότητα $P(U^2 + V^2 \leq 1)$.

δ) Να υπολογίσετε την δεσμευμένη κατανομή της X δεδομένου ότι $V = 2$.

5. Έστω (X, Y) από κοινού κανονικές $\mathscr{N}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\right)$.

α) Να βρείτε την από κοινού κατανομή των τ.μ. $U := X + Y$ και $W := X - Y$.

β) Να βρείτε επίσης την δεσμευμένη κατανομή της U δεδομένου ότι $W = 3$.

6. Έστω (X, Y) από κοινού κανονικές με μηδενικούς μέσους και πίνακα διακύμανσης – συνδιακύμανσης $\begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}$.

α) Υπολογίστε την $E[Y|X = 2]$.

β) Να υπολογίσετε την πιθανότητα $P(X + Y > 1)$.

γ) Να υπολογίσετε τις πιθανότητες $P(X > Y)$ και $P(X > Y|X > 0, Y > 0)$.

δ) Να υπολογίσετε τις πιθανότητες $P(X > 1|Y = 1)$ και $P(X+Y > 1|X-Y = 1)$.

7. Οι βαθμοί των μαθητών σε δύο τεστ, ένα γλωσσικό και ένα μαθηματικό, στην κλίμακα 0–100, ακολουθούν διμεταβλητή κανονική κατανομή με μέσους $(62, 57)$ και πίνακα διακυμάνσεων-συνδιακυμάνσεων $\begin{pmatrix} 225 & 117 \\ 117 & 169 \end{pmatrix}$.

α) Αν η βάση είναι το 50, ποια είναι η πιθανότητα να περάσει ένας μαθητής τα μαθηματικά;

β) Ένας μαθητής που πήρε βαθμό 87 στη γλώσσα, ποιά είναι η πιθανότητα να πέρασε τα μαθηματικά;

8. Έστω ότι οι τα ζεύγη των τυχαίων μεταβλητών $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, είναι ανεξάρτητα. Κάθε ζεύγος έχει διμεταβλητή κανονική κατανομή $\mathscr{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1/3 \\ 1/3 & 1 \end{pmatrix}\right)$.

α) Ποια είναι η πιθανότητα $P(X_1 + X_2 + X_3 > 2Y_1)$;

β) Ποια είναι η πιθανότητα $P(X_1 + 2X_2 > Y_1 - 3Y_2 + 1)$;

γ) Να υπολογίσετε την $E[X_1 + X_2 | Y_1 = 1]$ και την $E[X_1 | Y_1 + Y_2 = 2]$.

δ) Να υπολογίσετε την $P(X_1 + 2X_2 + 3X_3 + 4X_4 > 5 | Y_1 = -1, Y_2 = 2, Y_3 = 1)$.

9. Έστω X, Y ανεξάρτητες, τυποποιημένες κανονικές τ.μ. Να βρείτε την από κοινού κατανομή των τ.μ. $U := X + 2Y$ και $W := 2X - Y$.

10. Έστω (X, Y) από κοινού κανονικές με μέσους $(1, 2)$ και πίνακα διακύμανσης – συνδιακύμανσης $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. **α)** Υπολογίστε την $E[Y|X = 3]$ συναρτήσει του ρ. **β)** Αν $\rho = 0.5$ ποια είναι η πιθανότητα $P(X + 2Y > 1)$;

11. Έστω ότι το ύψος (σε cm) των ατόμων ενός πληθυσμού είναι κανονικά κατανεμημένο με μέσο 174 και τυπική απόκλιση 8. Διαλέγουμε με τυχαία δειγματοληψία 100 άτομα από τον πληθυσμό.

α) Ποια είναι η πιθανότητα ο αριθμητικός μέσος των υψών των 100 ατόμων του δείγματος να υπερβαίνει τα 175 cm;

β) Ποια είναι η πιθανότητα το ύψος του ψηλότερου ατόμου του δείγματος των 100 ατόμων να υπερβαίνει τα 192 cm;

γ) Υποθέτουμε τώρα ότι το ύψος των ατόμων του πληθυσμού είναι κατανεμημένο σύμφωνα με κάποια άλλη (μη κανονική) κατανομή με τον ίδιο μέσο (174 cm) και την ίδια τυπική απόκλιση (8 cm). Ισχύει (τουλάχιστον προσεγγιστικά) κάποια από τις απαντήσεις στα προηγούμενα δύο ερωτήματα και αν ναι ποια ή ποιες και γιατί;

12. Έστω (X, Y) από κοινού κανονικές με μέσους $(1, 2)$ και πίνακα διακύμανσης – συνδιακύμανσης $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. **α)** Να υπολογίσετε την δεσμευμένη πιθανότητα $P(Y > 3|X = 2)$. **β)** Ποια είναι η κατανομή της τυχαίας μεταβλητής $Z = 2X + 3Y$;

13. Το ύψος (σε cm) και το βάρος (σε kg) των ατόμων ενός πληθυσμού είναι από κοινού κανονικά κατανεμημένο με μέσους $(175, 75)$ και πίνακα διακύμανσης – συνδιακύμανσης $\begin{pmatrix} 64 & 76.8 \\ 76.8 & 144 \end{pmatrix}$.

α) Διαλέγουμε στην τύχη ένα άτομο από τον πληθυσμό. Ποια είναι η πιθανότητα το ύψος του να υπερβαίνει τα 180 cm;

β) Διαλέγουμε στην τύχη τρία άτομα από τον πληθυσμό και μετράμε τα ύψη τους. Δεδομένου ότι είναι 170, 180 και 190 cm αντίστοιχα ποια είναι η πιθανότητα το συνολικό τους βάρος να υπερβαίνει τα 250 kg;

14. Έστω X, Y, ανεξάρτητες, κανονικές, τυποποιημένες τ.μ. Να υπολογίσετε τις πιθανότητες

α) $P(|X| \leq 1)$,

β) $P(|X| \leq 1, |Y| \leq 1)$,

γ) $P(|X + Y| \leq 1)$,

δ) $P(|X + Y| \leq 1, |X - Y| \leq 1)$.

Πίνακας Τυποποιημένης Κανονικής Κατανομής

$$P(0 \leq Z \leq \alpha)$$

α	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990