הרצאה 7: משתנה מקרי רציף

בהרצאה זו נתבונן במשתנה מקרי רציף על מנת להבין מה ההבדל בינו למשתנה מקרי בדיד. גם נלמד מספר התפלגויות מיוחדות עבור משתנה מקרי רציף.

7.1. הגדרות בסיסיות

7.1.1. משתנה מקרי רציף

משתנה מקרי בדיד הוא משתנה אשר יכול לקבל רק מספר סופי (או בן מנייה) של ערכים בכל רווח נתון של ערכים ממשיים. למשל: מספר בנים במשפחה, מספר אנשים שפנו בין שעות מסוימות אל פקיד פלוני, וכדומה.

לעומת זאת, <u>משתנה מקרי רציף</u> (continuous random variable) הוא משתנה אשר יכול לקבל <u>כל ערך</u> בתוך רווח מספרים נתון. למשל: גבוהו של גבר מבוגר (כל מספר בין 150 ל-220 ס"מ); אורך חיים של אוטו; זמן המתנה אצל פקיד הדואר וכו'.

7.1.2. למה אנחנו צריכים להגדיר פונקצית צפיפות?

למדנו כי ניתן לתאר <u>משתנה מקרי בדיד</u> על ידי פונקצית הסתברות $P_X(x)$ אשר מביאה הסתברות <u>סופית</u> $P(X=x_i)$ למצוא $X=x_i$ כאשר x_i הוא ערך אפשרי ספציפי של מ.מ. X. למשל, אם X הוא מספר שמופיע על פאה בהטלת קובייה מאוזנת, $P(X=5)=\frac{1}{6}$.

זה לא המצב עבור מ.מ. רציף. ניתן לראות כי לכל תוצאה אפשרית של <u>משתנה מקרי רציף</u> מתאימה <u>הסתברות אפס</u>.

כדי להבין את הטענה הנ"ל, באו נזרוק מבט על מחוג גדול של שעון. הסתברות למצוא את המחוג בין הזוויות α ל- β שווה ל-

$$P(\alpha \le \theta \le \beta)=\frac{\beta-\alpha}{360}.$$

רואים כי הסתברות $P(\alpha \le \theta \le \beta)$ למצוא את המשתנה המקרי θ בתווך מסוים (בין α ל- β) היא סופית. למרות זאת, ההסתברות למצוא את המחוג מצביע על זווית מסוימת שווה בדיוק ל-0 מכיוון ש-

$$P(\theta=\alpha)=\lim_{\beta\to\alpha} P(\alpha \le \theta \le \beta)=\frac{\alpha-\alpha}{360}=0.$$

מהדוגמא עולה כי – עבור מ.מ. רציף – השאלה הבסיסית היא לא "מה ההסתברות שהמשתנה יקבל ערך מסוים?" אלא "מה ההסתברות שהמשתנה יקבל ערך <u>בתווך רווח מסוים</u>?"

חשוב! בגלל הטענה $P(\theta=\alpha)=0$ ו-$P(\theta=\beta)=0$ אין הבדל בין $P(\alpha\le\theta\le\beta)$ ל-$P(\alpha<\theta<\beta)$.

התאפסות של $P(\theta=\alpha)=0$ מביאה לצורך להגדיר תחליף לפונקצית הסתברות כאשר מדובר על מ.מ. רציף. תחליף כזה הוא <u>פונקצית צפיפות</u>.

<u>הגדרה</u>: <u>**פונקצית הצפיפות**</u> (density function) של משתנה מקרי רציף X היא פונקציה ממשית $f_X(x)$ אשר מוגדרת עבור $-\infty<x<\infty$ מקיימת בתחום זה את התכונות הבאות:

א. $f_X(x)\ge 0$

ב. לכל שני מספרים ממשיים a ו-b המקיימים $b>a$, ההסתברות $P(a<X<b)$ שווה לשטח מתחת לעקומה $y=f_X(x)$ בין a לבין b כך ש-

$$P(a<X<b)=\int_a^b f(x)dx.$$

ג. השטח הכולל בין $y=f(x)$ לבין ציר X שווה ל-1: $\int_{-\infty}^{+\infty} f(x)dx=1$ (נירמול).

7.1.3. פונקצית התפלגות מצטברת והיחס בינה לפונקצית צפיפות

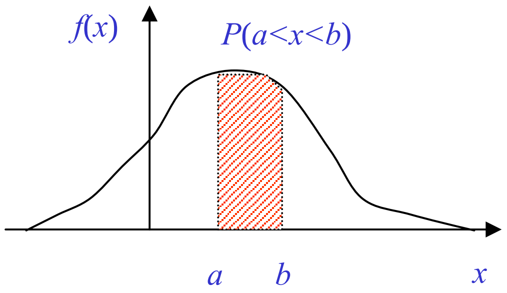
<u>תזכורת</u>: עבור מ.מ. X כלשהו, פונקצית התפלגות מצטברת (cumulative distribution function) שווה ל- $F_X(t)=P(X\le t)$.

מהגדרה זו ומהסעיף הקודם נובע כי $F_X(t)=\int_{-\infty}^{t} f_X(x)dx$.

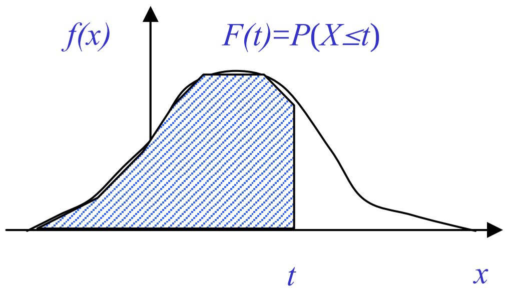
מזה ניתן לראות כי מתקיים $f_X(x)=\frac{d}{dx}F_X(x)$.

7.1.4. גרפים: הסתברות כשטח

<u>**הסתברות**</u>



<u>**פונקצית התפלגות מצטברת**</u>



<u>שאלה 1</u>: פונקצית הצפיפות של משתנה מקרי X נתונה על ידי הנוסחא

$$f_X(x)=\begin{cases}1-\frac{x}{2}, & 0\le x\le 2\\ 0, & \text{אחרת}\end{cases}$$

א. שרטט את הפונקציה, ודא שאמנם זו פונקצית צפיפות.

ב. חשב ושרטט את $F_X(t)$.

ג. חשב את ההסתברויות הבאות: $P(0.3<X<1.8), P(X\ge 1.5), P(X<0.2)$.

7.1.5. תוחלת ושונות של מ.מ. רציף

א. **<u>תוחלת</u>** של משתנה מקרי רציף ניתנת על ידי הביטוי $E[X]=\int_{-\infty}^{+\infty} xf_X(x)dx$. הנוסחא הזאת מקבילה לביטוי $E[X]=\sum_x x\ P_X(x)$ עבור משתנה מקרי בדיד.

ב. **<u>שונות</u>** של משתנה מקרי רציף ניתנת על ידי הביטוי

$$Var[X]=E[X^2]-(E[X])^2=\int_{-\infty}^{+\infty} x^2 f_X(x)dx-\left(\int_{-\infty}^{+\infty} xf_X(x)dx\right)^2$$

אשר מקביל לביטוי $Var[X]=\sum_x x^2 P_X(x)-\left(\sum_x x\ P_X(x)\right)^2$ עבור משתנה מקרי בדיד.

ג. **<u>סטיית תקן</u>** של משתנה מקרי רציף X שווה ל- $\sigma=\sqrt{Var[X]}$.

7.2. התפלגויות מיוחדות רציפות

7.2.1. התפלגות אחידה

<u>הגדרה</u>: משתנה מקרי רציף X נקרא **<u>מ.מ. אחיד</u>** בעל פרמטרים a ו- b, אם הוא מתואר על ידי פונקצית הצפיפות

$$f_X(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0 & \text{אחרת}\end{cases}$$

<u>סימון</u>: $X\sim U(a,b)$.

<u>מודל מתמטי</u>: משתנה מקרי רציף אחיד קשור למודל של <u>בחירה מקרית של מספר ממשי</u> שערכיו בין a ל- b.

<u>דוגמא</u>: מחוג של שעון מהסעיף 6.1.2. ניתן לראות כי זווית עליה מצביע המחוג מתפלגת התפלגות אחידה עם הפרמטרים $a=0$ ו- $b=360$ כך ש-

$$f_\Theta(\theta)=\begin{cases}0, & \theta<0\\ 1/360, & 0<\theta<360\\ 0, & \theta>360\end{cases}\qquad F_\Theta(t)=\begin{cases}0, & t<0\\ \dfrac{t}{360}, & 0\le t<360\\ 1, & t\ge 360\end{cases}$$

<u>תוחלת של משתנה אחיד:</u> $E[X]=\int_{-\infty}^{+\infty}xf_X(x)dx=\frac{1}{b-a}\int_a^b xdx=\frac{a+b}{2}$.

<u>שונות של משתנה אחיד:</u> $Var[X]=\frac{1}{b-a}\int_a^b x^2dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}$.

<u>פונקצית התפלגות מצטברת:</u> $F_X(t)=\int_{-\infty}^{t}f_X(x)dx=\begin{cases}0, & t<a\\ \dfrac{t-a}{b-a}, & a\le t<b\\ 1, & t\ge b\end{cases}$.

<u>שאלה 2</u>

במשוואה ריבועית $x^2-ax+b=0$ הפרמטרים a ו- b הם משתנים מקריים רציפים בלתי תלויים בעלי התפלגות אחידה: $a\sim U(0,1)$ ו- $b\sim U(0,1)$. נסמן את שורשיה ב- x_1 ו- x_2.

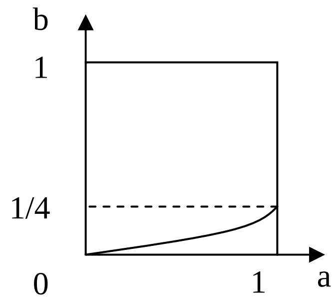
א. מה ההסתברות ששני השורשים ממשיים?

ב. אם ידוע שכל השורשים ממשיים, מה ההסתברות ש- $|x_1-x_2|<1/2$?

<u>פתרון:</u>

שורשי המשוואה הם $x_1=\frac{a-\sqrt{a^2-4b}}{2}$ ו- $x_1=\frac{a+\sqrt{a^2-4b}}{2}$.

א. שורשים הם ממשיים אם מתקיים אי שיווין $a^2-4b>0$ כך שההסתברות הדרושה היא $P(b<a^2/4)$ אשר מתאימה לשטח מתחת לעקומה $b=a^2/4$:



אזי $P(b<a^2/4)=\int_0^1\frac{a^2}{4}da=\frac{1}{12}$.

ב. באותה דרך, מכיוון ש- $|x_1 - x_2| = \sqrt{a^2 - 4b}$, ההסתברות הדרושה היא

$$P\left(\sqrt{a^2-4b} < \frac{1}{2} / a^2 > 4b\right) = P\left(a^2 - 4b < \frac{1}{4} / a^2 > 4b\right)$$

$$= \frac{P\left(\left\{a^2 - 4b < \frac{1}{4}\right\} \cap \left\{a^2 > 4b\right\}\right)}{P\left(a^2 > 4b\right)} = \frac{P\left(\frac{a^2}{4} - \frac{1}{16} < b < \frac{a^2}{4}\right)}{P\left(b < \frac{a^2}{4}\right)}$$

ניתן לראות כי זה שווה ל- $\frac{1}{2} = \frac{\int_0^1 \frac{a^2}{4} da - \int_{1/2}^1 \left(\frac{a^2}{4} - \frac{1}{16}\right) da}{\frac{1}{12}}$.

7.2.2. התפלגות מערכית

<u>הגדרה</u>: משתנה מקרי רציף Y נקרא <u>מ.מ. מערכי</u> בעל פרמטר λ אם הוא מתואר על ידי פונקצית הצפיפות

$$f_Y(y) = \begin{cases} \lambda \exp(-\lambda y), & y \ge 0 \\ 0, & y < 0 \end{cases}$$

<u>סימון</u>: $Y \sim Exp(\lambda)$.

<u>מודלים מתמטיים המביאים למשתנה מערכי</u>:

א. משך הזמן עד התרחשות האירוע הראשון בזרם אירועים פואסוני.

ב. אורך חיי מכשיר חשמלי או מכני.

<u>הסברים והוכחות</u>:

נתבונן בזרם אירועים פואסוני בעל קצב λ (השווה לממוצע של התרחשויות ביחידת זמן). אם נגדיר משתנה מקרי <u>בדיד</u> $X = \{$מספר אירועים במשך יחידת זמן$\}$, המשתנה הזה יתפלג פואסונית: $X \sim P(\lambda)$ כך שההסתברות

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

עבור $k = 0,1,2,...$.

אם נתבונן בפרק זמן $(0,t)$, מספר האירועים X_t המתרחשים במשך זמן זה יתפלג פואסונית אם פרמטר λt : $X_t \sim P(\lambda t)$ כך שההסתברות

$$P(X_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

עבור $k = 0,1,2,...$.

באו נגדיר משתנה **רציף** $Y=\{$משך הזמן שיחלוף עד התרחשותו האירוע הראשון$\}$. ברור כי המשתנה Y יכול לקבל כל ערך בתווך $(0,\infty)$. **איך מתפלג ה-Y ?**

כדי לענות על השאלה, נצטרך לחשב את ההסתברות $P(Y>t)$ שעד התרחשותו של האירוע הראשון יחלוף זמן גדול מ-t. ההסתברות $P(Y>t)=P(X_t=0)$ כך ש-

$$P(Y>t)=\frac{(\lambda t)^0}{0!}e^{-\lambda t}=\exp(-\lambda t).$$

מזה נובע כי $P(Y<t)=1-\exp(-\lambda t)$.

דרך זו אנחנו מגיעים למסכנה ש<u>פונקצית התפלגות מצטברת</u> עבור המשתנה המקרי Y ניתנת על ידי הנוסחא

$$F_Y(t)=P(Y<t)=1-\exp(-\lambda t).$$

אם ניקח בחשבון קשר בין פונקצית התפלגות מצטברת לפונקצית צפיפות – $f_Y(y)=\frac{d}{dy}F_Y(y)$ – נקבל:

$$f_Y(y)=\lambda\exp(-\lambda y)$$

עבור $y\geq 0$. **<u>זה מוכיח כי משך הזמן עד התרחשותו האירוע הראשון מתפלג מערכית עם פרמטר λ</u>** (השווה לממוצע של התרחשויות ביחידת זמן).

<u>תוחלת של משתנה מערכי:</u> $E[Y]=\int_{-\infty}^{+\infty}yf_Y(y)dy=\lambda\int_0^{\infty}y\exp(-\lambda y)dy=\frac{1}{\lambda}$.

<u>שונות של משתנה מערכי:</u> $Var[Y]=\lambda\int_0^{\infty}y^2\exp(-\lambda y)dy-\frac{1}{\lambda^2}=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$.

<u>שאלה 3</u>

נא להוכיח **<u>תכונת "חוסר זיכרון"</u>**: $P(Y>s+t\ /\ Y>s)=P(Y>t)$.

<u>פתרון:</u>

$$P(Y>s+t\ /\ Y>s)=\frac{P(Y>s+t\ \cap\ Y>s)}{P(Y>s)}=\frac{P(Y>s+t)}{P(Y>s)}$$
$$=\frac{\exp[-\lambda(s+t)]}{\exp[-\lambda s]}=\exp(-\lambda t)=P(Y>t)$$

<u>שאלה 4</u>

ניתן להניח שהבקעות השערים במשחק כדורגל מהווה זרם אירועים פואסוני. בדיקה העלתה שבמשחקי הליגה הלאומית מובקעים בממוצע 3 שערים במשחק (90 דקות). השתמש בהתפלגות המערכית כדי לחשב את ההסתברויות הבאות:

א. ההסתברות שהשער הראשון במשחק יובקע רק במחצית השנייה.

ב. נכנסת למשחק באיחור והתוצאה הייתה 1:0. מהי ההסתברות שבחצי השעה הבאה התוצאה לא תשתנה?

פתרון: הפרמטר λ של זרם אירועים פואסוני שווה לממוצע של הבקעות ביחידת זמן (= דקה 1). אזי, $\lambda=\frac{3}{90}=\frac{1}{30}$.

א. נגדיר משתנה מקרי Y = {זמן המתנה עד הבקעת השער הראשון}. ברור שהוא מתפלג מערכית. בגלל זה, ההסתברות שהשער הראשון במשחק יובקע רק במחצית השנייה שווה ל- $P(Y>45)$.

ב. בגלל תכונת "חוסר זיכרון", ההסתברות הדרושה היא $P(Y>30)$.

7.2.3. התפלגות נורמלית

<u>הגדרה</u>: משתנה מקרי רציף X נקרא <u>מ.מ. נורמלי</u> אם הוא מתואר על ידי פונקציית הצפיפות

$$f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

עבור $-\infty<x<+\infty$.

<u>סימון</u>: $X\sim N(\mu,\sigma^2)$

כדי להבין מהי המשמעות של הפרמטרים μ ו- σ^2 יש לחשב את תוחלתו ושונותו של משתנה מקרי $X\sim N(\mu,\sigma^2)$.

<u>תוחלת של משתנה נורמלי:</u> $E[X]=\int_{-\infty}^{+\infty}xf_X(x)dx=\mu$.

<u>שונות של משתנה נורמלי:</u> $Var[X]=\int_{-\infty}^{+\infty}x^2f_X(x)dx-\mu^2=\sigma^2$.

<u>הגדרה</u>: משתנה נורמלי Z בעל תוחלת 0 ושונות 1 נקרא <u>משתנה נורמלי סטנדרטי</u> – $Z\sim N(0,1)$ (standard normal variable). פונקציית צפיפות עבורו ניתנת על ידי הנוסחא:

$$f_Z(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

עבור $-\infty<z<+\infty$.

<u>טענה</u>: אם $X\sim N(\mu,\sigma^2)$, המשתנה הנורמלי הסטנדרטי מתקבל על ידי הנוסחא $Z=\frac{X-\mu}{\sigma}$.

<u>שאלה 5</u>
נא להוכיח את הטענה הנ"ל באמצעות תכונות של תוחלת ושונות.

<u>פונקצית התפלגות מצטברת</u> של משתנה נורמלי סטנדרטי שווה ל-

$$P(Z \le t) = \int_{-\infty}^{t} f_Z(z)dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{z^2}{2}}dz = \Phi(t).$$

יש להתיחס לפונקצית- Φ כמו לפונקציה חדשה.

<u>שאלה 6</u>

נא להוכיח כי $\Phi(-t) = 1 - \Phi(t)$

א. דרך גראפית,

ב. דרך נוסחאות.

<u>שאלה 7</u>

חשב $\Phi(0.91), \Phi(2.03), \Phi(0.02), \Phi(3.06), \Phi(0.41), \Phi(1.23)$ באמצעות הטבלה של פונקצית- Φ.

<u>שאלה 8</u>

נא להוכיח כי עבור $X \sim N(\mu, \sigma^2)$ מתקיים:

א. $P(X \le a) = \Phi\left(\frac{a-\mu}{\sigma}\right)$,

ב. $P(a \le X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$.

<u>שאלה 9</u>

חשב את ההסתברות $P(5 \le X \le 10)$ עבור $X \sim N(8,4)$.

<u>שאלה 10</u>

נא להוכיח כי

א. $P(\mu - \sigma \le X \le \mu + \sigma) \approx 0.68$,

ב. $P(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 0.95$.

ג. תדגים את התכונות הנ"ל באמצעות הציור ותן ניסוח כמותי למשמעות של סטיית התקן של מ.מ. נורמלי.

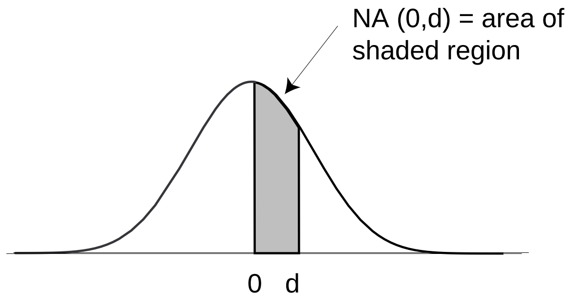
<u>תכונה חשובה של התפלגות נורמלית:</u> סכום של שני משתנים נורמליים <u>בלתי תלויים</u> הוא משתנה נורמלי גם כן!

<u>שאלה 11</u>

נתונים שני משתנים נורמליים בלתי תלויים $X_1 \sim N(\mu_1, \sigma_1^2)$ ו- $X_2 \sim N(\mu_2, \sigma_2^2)$. הוכח כי סכומם $X = X_1 + X_2$ מתפלג נורמלית: $X \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Appendix to the Lecture No 7

Normal distribution table: Don't forget to add 1/2 !!



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000