

Topics covered in this tutorial

- Moment-generating functions, sums and moments
- Poisson sums and Gaussian approximations
- The central limit theorem and sample means
- Probability bounds and a random-vector probability calculation

Complete all questions in this tutorial before the tutorial test that follows. The test normally starts between approximately 16:00 and 16:15.

Onderwerpe wat in hierdie tutoriaal gedek word

- Momentgenererende funksies, somme en momente
- Poisson-somme en Gaussiese benaderings
- Die sentrale limietstelling en steekproefgemiddeldes
- Waarskynlikheidsgrense en 'n toevalsvektor-waarskynlikheidsberekening

Voltooi al die vrae in hierdie tutoriaal voor die tutoriaaltoets wat daarop volg. Die toets begin gewoonlik tussen ongeveer 16:00 en 16:15.

1. Random variables J and K have the joint probability mass function

$P_{J,K}(j,k)$	$k = -1$	$k = 0$	$k = 1$
$j = -2$	0.35	0.15	0.10
$j = -1$	0.21	0.09	0.10

- (a) What is the moment generating function of J ? Wat is die moment genererende funksie van J ?
 - (b) What is the moment generating function of K ? Wat is die moment genererende funksie van K ?
 - (c) What is the probability mass function of $M = J + K$? Wat is die waarskynlikheidsmassafunksie van $M = J + K$?
 - (d) What is $E[M^2]$? (Use the MGF) Wat is $E[M^2]$? (Gebruik die MGF)
2. At time $t = 0$, you begin counting the arrivals of buses at a bus station. The number of buses K_i that arrive between time $i - 1$ minutes and time i minutes, has the Poisson PMF

$$P_{K_i}(k) = \begin{cases} 2^k e^{-2}/k! & k = 0, 1, 2, \dots, \\ 0 & \text{otherwise} \end{cases}$$

and $K_1, K_2, \dots$ are an independent and identically distributed random sequence. Let $R_i = K_1 + K_2 + \dots + K_i$ denote the number of buses arriving in the first i minutes.

en $K_1, K_2, \dots$ is 'n onafhanklike en identies verspreide stogastiese ry. Laat $R_i = K_1 + K_2 + \dots + K_i$ die hoeveelheid busse wees wat arriveer in die eerste i minute.

- (a) What is the moment generating function $\phi_{K_i}(s)$? Wat is die momentgenererende funksie $\phi_{K_i}(s)$?
- (b) Find the MGF $\phi_{R_i}(s)$. Vind die MGF $\phi_{R_i}(s)$.
- (c) Find the PMF $P_{R_i}(r)$. Hint: Compare $\phi_{R_i}(s)$ and $\phi_{K_i}(s)$. Vind die WMF $P_{R_i}(r)$. Wenk: Vergelyk $\phi_{R_i}(s)$ en $\phi_{K_i}(s)$.
- (d) Find $E[R_i]$ and $\text{Var}[R_i]$. Vind $E[R_i]$ en $\text{Var}[R_i]$.

3. Three statistically independent random variables X_1 , X_2 and X_3 with unknown distributions have the following means and variances: Drie statisties onafhanklike toevalsveranderlikes X_1 , X_2 en X_3 met onbekende verdelings het die volgende gemiddeldes en variansies:

$$\begin{aligned}\mu_{X_1} &= 2.5 & \sigma_{X_1}^2 &= 1.8 \\ \mu_{X_2} &= -0.6 & \sigma_{X_2}^2 &= 2.4 \\ \mu_{X_3} &= 1.2 & \sigma_{X_3}^2 &= 0.9\end{aligned}$$

- (a) The distributions of X_1 , X_2 and X_3 are unknown; only their means, variances and independence are given. Because no further distributional information is available, use the central limit theorem (CLT) to construct a Gaussian approximation for the PDF of $X = X_1 + X_2 + X_3$, using the mean and variance of X . Show your working. Does the information given guarantee that this approximation is accurate? Explain briefly. Die verdelings van X_1 , X_2 en X_3 is onbekend; slegs hul gemiddeldes, variansies en onafhanklikheid word gegee. Omdat geen verdere inligting oor die verdelings beskikbaar is nie, gebruik die sentrale limietstelling (SLT) om 'n Gaussiese benadering vir die WDF van $X = X_1 + X_2 + X_3$ te konstrueer, met gebruik van die gemiddelde en variansie van X . Wys u bewerkings. Waarborg die gegewe inligting dat hierdie benadering akkuraat is? Verduidelik kortliks.
- (b) Using the Gaussian model in (a), estimate $P(X \geq 5.36)$. Gebruik die Gaussiese model in (a) om $P(X \geq 5.36)$ te beraam.
4. $X_1, \dots, X_n$ is an iid sequence of exponential random variables, each with expected value 8. Let $M_n(X) = \frac{1}{n} \sum_{i=1}^n X_i$ be the sample mean. $X_1, \dots, X_n$ is 'n oiv ry eksponensiële toevalsveranderlikes, elk met 'n verwagte waarde van 8. Laat $M_n(X) = \frac{1}{n} \sum_{i=1}^n X_i$ die steekproefgemiddelde wees.
- (a) What is $\text{Var}[M_{16}(X)]$, the variance of the sample mean based on sixteen trials? Wat is $\text{Var}[M_{16}(X)]$, die variansie van die steekproefgemiddelde gebaseer op sestien steekproewe?
- (b) What is $P(X_1 > 10)$, the probability that one outcome exceeds 10? Wat is $P(X_1 > 10)$, die waarskynlikheid dat een uitkoms groter as 10 is?
- (c) Estimate $P(M_{16}(X) > 10)$, the probability that the sample mean of sixteen trials exceeds 10. (**Hint:** Use the central limit theorem.) Beraam $P(M_{16}(X) > 10)$, die waarskynlikheid dat die steekproefgemiddelde van 16 waarnemings groter as 10 is. (**Wenk:** Gebruik die sentrale limietstelling.)
5. Let $X_1, X_2, \dots$ be independent and identically distributed (iid) copies of X , with finite variance $\text{Var}[X]$. Define a new random sequence by Laat $X_1, X_2, \dots$ onafhanklike en identies verdeelde (oiv) kopieë van X wees, met eindige variansie $\text{Var}[X]$. Definieer 'n nuwe toevalsry deur

$$Y_n = X_{2n-1} - X_{2n}, \quad n = 1, 2, \dots$$

- (a) What is $E[Y_n]$? Wat is $E[Y_n]$?
- (b) What is $\text{Var}[Y_n]$? Wat is $\text{Var}[Y_n]$?
- (c) What is the expected value and variance of $M_n(Y)$ in terms of X ? Wat is die verwagte waarde en variansie van $M_n(Y)$ in terme van X ?

6. Two fair, independent dice are rolled together repeatedly, with independent results between rolls. A success (“snake eyes”) occurs when both dice show one. Let R be the number of rolls of the pair, including the final roll, needed to obtain the third success. Find
- Twee billike, onafhanklike dobbelstene word herhaaldelik saam gerol, met onafhanklike uitkomstes tussen rolle. ’n Sukses (“snake eyes”) vind plaas wanneer albei een wys. Laat R die aantal rolle van die paar wees, met die laaste rol ingesluit, wat nodig is vir die derde sukses. Vind
- (a) the upper bound to $P(R \geq 250)$ based on the Markov inequality, die boonste grens tot $P(R \geq 250)$ baseer op die Markov ongelykheid,
- (b) the upper bound to $P(R \geq 250)$ based on the Chebyshev inequality, die boonste grens tot $P(R \geq 250)$ baseer op die Chebyshev ongelykheid,
- (c) the exact value of $P(R \geq 250)$. Hint: Count successes in the first 249 rolls. die presiese waarde van $P(R \geq 250)$. Wenk: Tel suksesse in die eerste 249 rolle.
7. Let $\mathbf{X}$ be a 2×1 continuous random vector and denote its components by X_1 and X_2 . Let the support of $\mathbf{X}$ be
- Laat $\mathbf{X}$ ’n 2×1 kontinue toevalsvektor wees en dui sy komponente aan met X_1 en X_2 . Laat die steunversameling van $\mathbf{X}$

$$S_{\mathbf{X}} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, 0 \leq x_2 \leq 3\}$$

that is, the set of all 2×1 vectors such that the first component belongs to the interval $[0, \infty)$ and the second component belongs to the interval $[0, 3]$. Let the joint probability density function of $\mathbf{X}$ be

wees, d.w.s. die stel van alle 2×1 vektore waarvan die eerste komponent aan die interval $[0, \infty)$ en die tweede komponent aan die interval $[0, 3]$ behoort. Laat die gesamentlike waarskynlikheidsdigtheidsfunksie van $\mathbf{X}$ gegee word deur

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1, X_2}(x_1, x_2) = \begin{cases} \frac{1}{3}e^{-x_1} & \text{if } \mathbf{x} \in S_{\mathbf{X}} \\ 0 & \text{otherwise} \end{cases}$$

Compute $P(X_1 + X_2 \leq 4)$

Bereken $P(X_1 + X_2 \leq 4)$