2. Review: Cumulative Distribution Functions

Warmup: Integration limits of CDF

3/3 points (graded)

The cumulative distribution function $\Phi : \mathbb{R} \to \mathbb{R}$ of the standard normal $\mathcal{N}(0,1)$ can be written as

$$\Phi(z) = \int_A^{B(z)} \frac{1}{\sqrt{2\pi}} e^{C(x)}\, dx$$

where $B(z)$ is a function of z and $C(x)$ is a function of x. Write down the integration limits $A, B(z)$, as well as the function $C(x)$ in the integrand.

Enter inf for ∞.

$A =$ -inf ✔ **Answer:** -inf

$-inf$

$B =$ z ✔ **Answer:** z

z

$C =$ -x^2/2 ✔ **Answer:** -x^2/2

$-\frac{x^2}{2}$

STANDARD NOTATION

Solution:

Recall the cdf of a distribution $\mathbf{P}$ is a function F such that

$$F(x) = P(X \le x), \quad X \sim \mathbf{P}.$$

The density of a standard Gaussian is given by $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Therefore, the CDF of $\mathcal{N}(0,1)$ is given by

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx.$$

Hence $A = -\infty$, $B = z$, and $C = -x^2/2$.

Submit You have used 3 of 3 attempts

Answers are displayed within the problem

Review

▶ 7:29 / 7:29 ▸ 1.0x

Now, of course, this is not exactly what you're testing,

right?

You're not testing if your data is standard Gaussian.

What you're testing is, is my data

likely to come from this histogram?

OK, from this pdf.

And this pdf is just the piecewise constant pdf

that's close to the Gaussian curve, but not exactly, OK?

End of transcript. Skip to the start.

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Review: Cumulative Distribution Function

0/1 point (graded)

Let X and Y be real-valued random variables, both distributed according to a distribution $\mathbf{P}$. (We make no assumption about their joint distribution). Let F denote the **cdf** of $\mathbf{P}$.

Which of the following are true about the cdf F? (Choose all that apply.)

- [x] $P(X \leq t)$ and $P(Y \leq t)$ are random variables.
- [x] For all $t \in \mathbb{R}$, $F(t) = P(X \leq t)$ and $F(t) = P(Y \leq t)$. ✔
- [ ] $F(t) = P(X \leq t) = P(Y \leq t)$ only if X and Y are independent.
- [ ] $\lim_{t \to \infty} F(t) = 1$. ✔
- [x] $\lim_{t \to -\infty} F(t) = 0$. ✔
- [ ] $\int_{-\infty}^{\infty} F(t)\, dt = 1$

Solution:

We examine the choices in order.

- The first choice is incorrect. The probability of an event is a real number, not a random variables, so both $\mathbf{P}(X \leq t)$ and $\mathbf{P}(Y \leq t)$ are deterministic numbers.
- The second choice is correct. The joint distribution of X and Y is irrelevant– so long as X and Y have the same distribution, it will be true that $\mathbf{P}(X \leq t) = \mathbf{P}(Y \leq t)$. By definition, both of these quantities are equal to $F(t)$.

- The third choice is incorrect. As stated in the previous bullet, regardless of the joint distribution of X and Y, as long as they are identically distributed, it is true that $\mathbf{P}(X \leq t) = \mathbf{P}(Y \leq t)$.

- The fourth choice is correct. Observe that $\boxed{\mathbf{P}(X \leq \infty)} = 1$ because X is real-valued. Moreover, F is an increasing function of t. Therefore, $\boxed{\lim_{t \to \infty} F(t) = \mathbf{P}(X \leq \infty) = 1.}$

- The fifth choice is correct, since $\lim_{t \to -\infty} F(t) = \mathbf{P}(X \leq -\infty) = 0$.

- The final choice is incorrect. The statement is true for the **pdf** not the cdf:

$$\int_{-\infty}^{\infty} f(t)\,dt = 1 \qquad \text{if } f(t) \text{ is pdf of } X.$$

Since the cdf $F(t)$ has limit $\lim_{t \to \infty} F(t) = 1$, the integral over $\mathbb{R}$ of $F(t)$ diverges.

Submit

You have used 2 of 2 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 4 Hypothesis testing:Lecture 16: Goodness of Fit Tests Continued: Kolmogorov-Smirnov test, Kolmogorov-Lilliefors test, Quantile-Quantile Plots / 2. Review: Cumulative Distribution Functions