توزیع آماری (Statistical Distribution) به مجموعه‌ای از داده‌ها یا مقادیر متغیری گفته می‌شود که بر اساس احتمال وقوع هر مقدار یا رخداد در یک بازه مشخص توزیع می‌شود. این توزیع‌ها به ما نشان می‌دهند که چگونه داده‌ها در طول یک بازه یا دامنه قرار دارند و چه مقداری با چه احتمالی رخ می‌دهد.

واریانس

**واریانس** معیاری از پراکندگی داده‌ها نسبت به میانگین است. این شاخص نشان می‌دهد که داده‌ها تا چه حد از میانگین فاصله دارند. به صورت محاسباتی، واریانس برابر است با میانگین مجذور تفاضل هر داده از میانگین کل داده‌ها.

به عبارت دیگر، واریانس نشان می‌دهد که داده‌ها چقدر پخش شده یا به میانگین نزدیک‌اند. واریانس بزرگ‌تر نشان‌دهنده پراکندگی بیشتر داده‌ها است و واریانس کوچک‌تر نشان‌دهنده تمرکز بیشتر داده‌ها به دور میانگین.

**انحراف معیار (Standard Deviation)**، اندازه‌ای از پراکندگی داده‌ها نسبت به میانگین است و به صورت جذر واریانس محاسبه می‌شود.

این شاخص نشان می‌دهد که داده‌ها به طور متوسط چقدر از میانگین فاصله دارند. انحراف معیار کوچک نشان‌دهنده تمرکز داده‌ها نزدیک به میانگین، و انحراف معیار بزرگ نشان‌دهنده پراکندگی بیشتر داده‌ها است.

توزیع‌های گسسته نوعی از توزیع‌های آماری هستند که برای متغیرهایی به کار می‌روند که فقط می‌توانند مقادیر گسسته (معمولاً اعداد صحیح) بگیرند. برخی از معروف‌ترین توزیع‌های گسسته عبارت‌اند از:

1. **توزیع دو جمله‌ای (Binomial Distribution)**: تعداد موفقیت‌ها در تعداد معینی از آزمایش‌های مستقل با دو نتیجه ممکن (مثلاً موفقیت و شکست) و احتمال ثابت برای موفقیت.
2. **توزیع پواسون (Poisson Distribution)**: احتمال تعداد رخدادهای یک رویداد خاص در یک بازه زمانی یا فضایی ثابت، زمانی که رخدادها مستقل هستند و با نرخ ثابت اتفاق می‌افتند.
3. **توزیع هندسی (Geometric Distribution)**: احتمال تعداد آزمایش‌هایی که برای دستیابی به اولین موفقیت در یک سری آزمایش‌های مستقل با دو نتیجه ممکن و احتمال ثابت برای موفقیت لازم است.
4. **توزیع چند جمله‌ای (Multinomial Distribution)**: گسترش توزیع دو جمله‌ای، برای زمانی که هر آزمایش می‌تواند چندین نتیجه ممکن داشته باشد (بیش از دو نتیجه).
5. **توزیع فوق‌هندسی (Hypergeometric Distribution)**: تعداد موفقیت‌ها در نمونه‌ای که بدون جایگذاری از یک جامعه با دو نوع عنصر مختلف گرفته شده است.

این توزیع‌ها اغلب در مدل‌سازی رویدادهایی به کار می‌روند که نتایجشان قابل شمارش و گسسته هستند.