Test 2 Summary

MH3500

October 16, 2024

1 Discrete Random Variables

(a) $X \sim$ Bernoulli(p).

- $P(X = x) = (1-p)^{1-x}p^x, \quad$ for $x = 0, 1$.
- $E(X) = p, \quad Var(X) = p(1-p), \quad M_X(t) = pe^t + 1 - p$

(b) $X \sim$ Binomial(n, p).

- $P(X = x) = \binom{n}{p}(1-p)^{1-x}p^x, \quad$ for $x = 0, 1, \ldots, n$.
- $E(X) = np, \quad Var(X) = np(1-p), \quad M_X(t) = (pe^t + 1 - p)^n$

(c) $X \sim$ Geometric(p).

- $P(X = x) = p(1-p)^{x-1}, \quad$ for $x = 1, 2, \ldots$.
- $E(X) = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}, \quad M_X(t) = \frac{pe^t}{1-(1-p)e^t}, \quad$ for $t < -\ln(1-p)$

(d) $X \sim$ Poisson(λ).

- $P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad$ for $x = 0, 1, \ldots$.
- $E(X) = \lambda, \quad Var(X) = \lambda, \quad M_X(t) = e^{\lambda(e^t-1)}$

2 Continuous Random Variables

(a) $X \sim$ Uniform(a, b).

- $f_X(x) = \frac{1}{b-a}, \quad$ for $x \in [a, b]$.
- $E(X) = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{12}, \quad M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}, \quad$ for $t \neq 0$

(b) $X \sim$ N(μ, σ^2).

- $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad$ for $x \in \mathbb{R}$.
- $E(X) = \mu, \quad Var(X) = \sigma^2, \quad M_X(t) = \exp\left(t\mu + \frac{\sigma^2 t^2}{2}\right)$

$M_X(t)$ is the moment generating function, please ignore for the time being.

3 Normal and Poisson Approximation to Binomial Distribution

Let $X \sim$Bin(n, p). If $n \geq 100, np \leq 10$, X can be approximated by a Poisson distribution with parameter $\lambda = np$. For example, if $n = 200, p = 0.02$. We have $X' \sim$Po(λ) to approximate X, where $\lambda = 200 \cdot 0.02 = 4$.

$$P(X = 100) \approx P(X' = 100) = e^{-4}\frac{4^{100}}{100!}.$$

Let $X \sim$Bin(n, p). If $np(1-p) \geq 10$, X can be approximated by a Normal distribution with parameter $\mu = np, \sigma^2 = np(1-p)$. For example, if $n = 1000, p = 0.05$. We have $X' \sim$N(μ, σ^2) to approximate X, where $\mu = 1000 \cdot 0.05 = 50$ and $\sigma^2 = 1000 \cdot 0.05 \cdot 0.95 = 47.5$.

$$P(a \leq X < b) \approx P(a - 0.5 \leq X' - 0.5) = P(\frac{a - 0.5 - \mu}{\sigma} \leq Z \leq \frac{b - 0.5 - \mu}{\sigma}), \quad Z \sim N(0, 1).$$

Remember the continuity correction when you approximate a discrete distribution using a continuous distribution.

$$P(X = k) \approx P(k - 0.5 \leq X' \leq k + 0.5)$$

$$\begin{aligned} P(X \leq k) &\approx P(X' \leq k + 0.5) \\ P(X \geq k) &\approx P(X' \geq k - 0.5) \\ P(X < k) &\approx P(X' \leq k - 0.5) \\ P(X > k) &\approx P(X' \geq k + 0.5) \end{aligned}$$

4 Other Useful Properties

- Suppose $X_1, \ldots X_n$ are i.i.d. random variables with a CDF F. Let $Y_1 = \min(X_1, \ldots, X_n)$ and $Y_2 = \max(X_1, \ldots, X_n)$, then
$$P(Y_1 \leq y) = 1 - (1 - F(y))^n \qquad P(Y_2 \leq y) = (F(y))^n.$$
- Finite geometric series:
$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r}.$$
- Infinite geometric series:
$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}, \quad |r| < 1.$$
- Binomial series:
$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a + b)^n.$$
- Taylor series expansion of e^x:
$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \quad \forall x \in \mathbb{R}.$$
- Gaussian Integral:
$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad \text{or} \quad \int_{0}^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$
  With a substitution $t = \frac{x-\mu}{\sqrt{2}\sigma}$, we can show that
$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1.$$