Probability Final

2017-04-24

- **Uniform distribution**: $f(x) = \frac{1}{b-a}$, $x \in [a, b]$
    - CDF: $F(x) = \frac{x-a}{b-a}$
    - MGF: $M(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}$
    - Median: $\frac{1}{2}(a+b)$
    - Mean: $\frac{1}{2}(a+b)$
    - Parameters: a, b
    - Variance: $\frac{1}{12}(b-a)^2$
    - Function in `Python`: `scipy.stats.uniform`
- **Exponential distribution**: $f(x) = \theta^{-1}\exp(-\frac{x}{\theta})$, $x \in [0, \infty)$
    - CDF: $F(x) = 1 - \exp(-\frac{x}{\theta})$
    - MGF: $M(t) = (1-\theta t)^{-1}$, for $t < \theta^{-1}$
    - Median: $\theta \ln(2)$
    - Mean: θ
    - Variance: θ^2
    - Function in `Python`: `scipy.stats.expon`
    - Parameters: θ (scale)
    - Literal: the time needed to observe the first occurrence.
    - Versus *geometric distribution*: the number of trials needed to observe the first occurrence.
    - The failure rate is constant.
    - A special case of *gamma distribution* whose $\alpha = 1$
- **Gamma distribution**: $f(x) = \Gamma(\alpha)^{-1}\theta^{-\alpha}x^{\alpha-1}\exp(-\frac{x}{\theta})$, $x \in [0, \infty)$
    - CDF: $F(x) = \Gamma(\alpha)^{-1}\gamma(\alpha, x/\theta)$
    - MGF: $M(t) = (1-\theta t)^{-\alpha}$ for $t < \theta^{-1}$
    - Median: -
    - Mean: $\alpha\theta$
    - Variance: $\alpha\theta^2$
    - Function in `Python`: `scipy.stats.gamma`
    - Parameters: α (shape), θ (scale)

- Literal: the time needed to observe the α-th occurrence.
- Versus *negative binomial distribution*: the number of trials needed to observe the α-th occurrence.
- **Gamma function**: $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$, also called **generalized factorial**. If α is a positive integer, then $\Gamma(\alpha) = (\alpha-1)!$
- **Lower incomplete gamma function**: $\gamma(\alpha, \beta) = \int_0^\beta t^{\alpha-1}e^{-t}dt$

- **Chi-square distribution**: $f(x) = \Gamma(\frac{d}{2})^{-1}2^{-\frac{d}{2}}x^{\frac{d}{2}-1}\exp(-\frac{x}{2})$, $x \in [0, \infty)$
  - CDF: $F(x) = \Gamma(\frac{d}{2})^{-1}\gamma(\frac{d}{2}, \frac{x}{2})$
  - MGF: $M(t) = (1-2t)^{-\frac{d}{2}}$, for $t < \frac{1}{2}$
  - Median: -
  - Mean: d
  - Variance: $2d$
  - Function in `Python`: `scipy.stats.chi2`
  - Parameters: d (degree of freedom)
  - A special case of *gamma distribution* whose $\alpha = \frac{d}{2}$, $\theta = 2$
- **Normal distribution**: $f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$. $x \in (-\infty, \infty)$
  - CDF: -
  - MGF: $M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$
  - Median: μ
  - Mean: μ
  - Variance: σ^2
  - Function in `Python`: `scipy.stats.norm`
  - Parameters: μ (mean), σ^2 (variance)

2017-05-08

- Relationship between normal and chi-square distribution: Given a normal distribution $X \sim N(\mu, \sigma^2)$. Then, $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, and $Z^2 = (\frac{X-\mu}{\sigma})^2 \sim \chi^2(1)$.
- Proof of $Z^2 = (\frac{X-\mu}{\sigma})^2 \sim \chi^2(1)$:
  $F(x) = P(Z^2 \le x) = P(|Z| \le \sqrt{x}) = 2 \cdot \int_{i=0}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2})\mathrm{d}z$
  $\therefore f(x) = F'(x) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})\frac{1}{\sqrt{x}} = \Gamma(\frac{1}{2})^{-1}\gamma(\frac{1}{2}, \frac{x}{2}) = \chi^2(1)$
- **Log normal distribution**: $f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp(-\frac{1}{2}(\frac{\ln x-\mu}{\sigma})^2)$. $x \in (0, \infty)$
  - CDF: $\Phi(\ln x - \mu)$
  - MGF: -

- Median: μ
- Mean: $\exp(\mu + \frac{1}{2}\sigma^2)$
- Variance: $(\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2)$
- Literal: $\ln(X) \sim N(\mu, \sigma^2)$

- Bivariate distribution of the discrete type:
  - **Joint probability mass function**: $f(x,y) = P(X = x, Y = y)$
  - **Marginal probability mass function** of X: $P(X = x) = f_X(x) = \sum_y f(x,y)$
  - **Marginal probability mass function** of Y: $P(Y = y) = f_Y(y) = \sum_x f(x,y)$
  - X and Y are **independent** random variables iff $P(X = x, Y = y) = P(X = x)P(Y = y)$
  - $E[u(X,Y)] = \sum\sum_{(x,y)\in S} u(x,y)f(x,y)$
  - **Covariance**: $\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X\mu_Y$
  - **Correlation coefficient**: $\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$
  - **Least squares regression line**: $y - \mu_Y = \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$
  - **Conditional probability mass function** of X: $f(x|y) = f(x,y)/f_Y(y)$
  - **Conditional probability mass function** of Y: $f(y|x) = f(x,y)/f_X(x)$
  - **Conditional mean** of X given $Y = y$: $\mu_{X|y} = E[X|y] = \sum_x xf(x|y)$
  - **Conditional mean** of Y given $X = x$: $\mu_{Y|x} = E[Y|x] = \sum_y yf(y|x)$
  - **Conditional variance** of X given $Y = y$:
    $\sigma^2_{X|y} = E[(X - \mu_{X|y})^2|y] = \sum_x (x - \mu_{X|y})^2 f(x|y) = E[X^2|y] - (\mu_{X|y})^2$
  - **Conditional variance** of Y given $X = x$:
    $\sigma^2_{Y|x} = E[(Y - \mu_{Y|x})^2|x] = \sum_y (y - \mu_{Y|x})^2 f(y|x) = E[Y^2|x] - (\mu_{Y|x})^2$

2017-05-15

- Bivariate distribution of the continuous type: the same concept as the discrete type.
- Bivariate normal distribution:
  - **Conditional mean** of Y given $X = x$: $\mu_{Y|x} - \mu_Y = \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$
  - **Conditional variance** of Y given $X = x$: $\sigma^2_{Y|x} = \sigma^2_Y(1 - \rho^2)$
- Independent v.s. Uncorrelated
  - For any distribution: independent → uncorrelated
  - For normal distribution: independent ↔ uncorrelated
- Multivariate normal distribution: $(2\pi)^{-\frac{1}{2}k}|\Sigma|^{-\frac{1}{2}}\exp(-\frac{1}{2}(x - \mu)^T\Sigma^{-1}(x - \mu))$
- [x] Homework: 4.1-8, 4.2-3, 4.3-7, 4.4-16, 4.5-10
- Solutions: [hw4_b00401062.pdf]()

2017-05-22

- Let X have a PDF that is $f(x)$, and $Y = u(x)$ be a function of X. The PDF of Y is $g(y) = f(u^{-1}(y))|(u^{-1})'(y)|$.
- Let Y have a distribution that is $U(0,1)$, and $F(x)$ be a cdf and strictly increasing on the support $a < x < b$. Then the random variable X defined by $X = F^{-1}(Y)$ is a continuous-type random variable with cdf $F(x)$.
- Let X have the cdf $F(x)$ of the continuous type that is strictly increasing on the support $a < x < b$. Then the random variable Y, defined by $Y = F(X)$, has a distribution that is $U(0,1)$.
- Given X_1 and X_2 have a joint PDF that is $f_X(x_1, x_2)$, and $Y_1 = u_1(X_1, X_2), Y_2 = u_2(X_1, X_2)$. The joint PDF of Y_1 and Y_2 is $f_Y(y_1, y_2) = f_X[u_1^{-1}(y_1, y_2), u_2^{-1}(y_1, y_2)]|J|$.
- Jacobian matrix J is $\begin{bmatrix} \frac{\partial u_1^{-1}}{\partial y_1} & \frac{\partial u_1^{-1}}{\partial y_2} \\ \frac{\partial u_2^{-1}}{\partial y_1} & \frac{\partial u_2^{-1}}{\partial y_2} \end{bmatrix}$
- The mean of a random sample $\overline{X} = \sum_{i=1}^{n} X_i/n$ from a distribution with mean μ and variance σ^2:
    - Mean of $\overline{X}$: μ
    - Variance of $\overline{X}$: σ^2/n
- If $X_1, X_2, \ldots, X_n$ are n observations of a random sample from a population, then
    - Sample mean: $\overline{X} = \sum_{i=1}^{n} X_i/n$
    - Sample variance: $S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/(n-1)$

2017-05-29

- **Laplace distribution**: $f(x;\theta) = \frac{1}{2}\theta^{-1}\exp(-\frac{|x|}{\theta}), x \in (-\infty, \infty)$
- **Beta distribution**: $f(x;\alpha,\beta) = \Gamma(\alpha+\beta)\Gamma(\alpha)^{-1}\Gamma(\beta)^{-1}x^{\alpha-1}(1-x)^{\beta-1} = \mathrm{B}(\alpha,\beta)x^{\alpha-1}(1-x)^{\beta-1}, x \in [0,1]$
- **F distribution**: $f(x;d_1,d_2) = \mathrm{B}(\frac{d_1}{2}, \frac{d_1}{2})^{-1}(\frac{d_1}{d_2})^{\frac{d_1}{2}}(1+\frac{d_1}{d_2}x)^{-\frac{d_1+d_2}{2}}x^{\frac{d_1}{2}-1}, x \in [0,\infty)$
- If $X_1, X_2, \ldots, X_n$ are independent random variables with respective moment-generating functions $M_{X_i}(t)$, $i = 1, 2, 3, \ldots, n$, then the moment-generating function of $Y = \sum_{i=1}^{n} c_i X_i$ is $M_Y(t) = \prod_{i=1}^{n} M_{X_i}(c_i t)$
- If X_1 and X_2 are two independent $\mathrm{Exponential}(\theta)$, then $X_1 - X_2 \sim \mathrm{Laplace}(\theta)$.
- If $X_1 \sim \mathrm{Gamma}(\alpha,\theta)$ and $X_2 \sim \mathrm{Gamma}(\beta,\theta)$ are independent, then $\frac{X_1}{X_1+X_2} \sim \mathrm{Beta}(\alpha,\beta)$
- If $X_1 \sim \chi^2(d_1)$ and $X_2 \sim \chi^2(d_2)$ are independent, then $\frac{X_1/d_1}{X_2/d_2} \sim F(d_1,d_2)$
- If $Z \sim N(0,1)$ and $U \sim \chi^2(d)$ are independent, then $\frac{Z}{\sqrt{U/d}} = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim T(d)$, where $d = n-1$

- If $X_1, X_2, \ldots, X_n$ are n independent $\text{Bernoulli}(p)$, then $\sum_{i=1}^n X_i \sim \text{Binomial}(n, p)$
- If $X_1, X_2, \ldots, X_n$ are n independent $\text{Geometric}(p)$, then $\sum_{i=1}^n X_i \sim \text{Negative Binomial}(n, p)$
- If $X_1, X_2, \ldots, X_n$ are n independent $\text{Exponential}(\theta)$, then $\sum_{i=1}^n X_i \sim \text{Gamma}(n, \theta)$
- If $X_1 \sim \chi^2(d_1)$, $X_2 \sim \chi^2(d_2)$, ..., $X_n \sim \chi^2(d_n)$ are independent, then $\sum_{i=1}^n X_i \sim \chi^2(\sum_{i=1}^n d_i)$
- If $Z_1, Z_2, \ldots, Z_n$ are n independent $N(0,1)$, then $\sum_{i=1}^n Z_i^2 \sim \chi^2(n)$
- If $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, ..., $X_n \sim N(\mu_n, \sigma_n^2)$ are independent, then $\sum_{i=1}^n c_i X_i \sim N(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2)$
- If $X_1, X_2, \ldots, X_n$ are n independent $N(\mu, \sigma^2)$, then $\overline{X} = \sum_{i=1}^n X_i / n \sim N(\mu, \sigma^2/n)$
- If $X_1, X_2, \ldots, X_n$ are n independent $N(\mu, \sigma^2)$, then $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$

2017-06-06

- **Central limit theorem**: $\overline{X} = \sum_{i=1}^n X_i / n$ is the mean of a random sample from a distribution with mean μ and variance σ^2. Then, $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- **Half-unit correction for continuity**: $P(Y = k) = P(k - \frac{1}{2} < Y < k + \frac{1}{2})$
- **Chebyshev's inequality**: $P(|X - \mu| \geq k\sigma) \leq 1/k^2$ or $P(|X - \mu| \geq \varepsilon) \leq \sigma^2/\varepsilon^2$
- **Law of large numbers**: $\lim_{n \to \infty} P(|\frac{Y}{n} - p| < \varepsilon) = 1$, i.e. $\frac{Y}{n}$ converges in probability to p when n is large enough.