W04

- **Definition**
  - Continous
    - Interval
    - P(X=c) = 0
  - Mix Discret and Continous
- **4.1 PDF**
  - $P(a \le X \le b) = \int_a^b f(x)dx$
  - Density curve
  - uniform distribution
    - $f(x; A, B) = \begin{cases} \frac{1}{B-A} & A \le x \le B \\ 0 & \text{otherwise} \end{cases}$
  - P(a <= X <= b) = P(a < X < b) = P(a < X <= b) = P(a <= X < b)
- **4.2 CDF and EV**
  - CDF
    - $F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)\,dy$
  - Using F(x) to Compute Probabilities
    - $P(a \le X \le b) = F(b) - F(a)$
  - Obtaining f(x) from F(x)
    - $F'(x) = f(x).$
  - Percentiles
    - $p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)\,dy$
    - Median
      - $.5 = F(\tilde{\mu}).$
  - EV
    - $\mu_X = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$
    - $E[h(X)] = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x) \cdot f(x)\,dx$
    - $\sigma_X^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)dx = E[(X-\mu)^2]$
    - $V(X) = E(X^2) - [E(X)]^2$
- **4.3 The Normal Distribution**
  - Central Limit Theorem
    - $f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \quad -\infty < x < \infty$
    - $X \sim N(\mu, \sigma^2).$
      - Mu: symmetric about mu and bell-shaped
      - Sigma: is the distance from mu to the inflection points of the curve
  - Standard normal distribution
    - $\mu = 0$ and $\sigma = 1$
    - $f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty$
    - cdf of Z is $P(Z \le z) = \int_{-\infty}^{z} f(y; 0, 1)\,dy$, denote by $\Phi(z)$.
  - Z_alpha for z Critical Values
    - z_α is the $100(1-\alpha)$th percentile of the standard normal distr
  - Nonstandard Normal Distributions
    - $Z = \frac{X-\mu}{\sigma}$
    - $P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$
    - $P(X \le a) = \Phi\left(\frac{a-\mu}{\sigma}\right) \qquad P(X \ge b) = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right)$
    - $(100p)\text{th percentile for normal } (\mu, \sigma) = \mu + \left[\begin{matrix}(100p)\text{th for} \\ \text{standard normal}\end{matrix}\right] \cdot \sigma$
  - SD
    - 1 SD: 68%
    - 2 SD: 95%
    - 3 SD: 99.7%
  - Approximating the binomial Distribution
    - $P(X \le x) = B(x; n, p) \approx \left(\begin{matrix}\text{area under the normal curve} \\ \text{to the left of } x + .5\end{matrix}\right) = \Phi\left(\frac{x + .5 - np}{\sqrt{npq}}\right)$
      - **0.5: continuity correction**
      - np >= 10
      - nq >= 10
- **Excel**
  - NORM.DIST — NORM.INV
  - GAMMA.DIST — GAMMA.INV — Verify each other!
  - EXPON.DIST
  - BETA.DIST
- **4.4 The exponential Distribution**
  - $f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$
    - $\mu = \frac{1}{\lambda} \qquad \sigma^2 = \frac{1}{\lambda^2}$
    - $F(x; \lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \ge 0 \end{cases}$
  - Poisson distribution, alpha t, P(x<t) cdf exponential distribution
    - $\lambda = \alpha.$
  - **"memoryless" property**
    - the distribution of additional lifetime is exactly the same as the original distribution of lifetime, => Not time dependent
- **4.4 Gamma family**
  - $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\,dx$
    1. For any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$
    2. For any positive integer, n, $\Gamma(n) = (n-1)!$
    3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
  - $f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$
    - **Alpha = 1, Beta = 1/Lambda: Exponential**
    - standard gamma distribution: Beta = 1 (Scale parameter)
    - Alpha <=1, pdf decrease
    - Alpha >1, pdf incease then decrease
    - $E(X) = \mu = \alpha\beta \qquad V(X) = \sigma^2 = \alpha\beta^2$
    - Beta = 1, standard gamma
      - $F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}\,dy \qquad x > 0$
    - $P(X \le x) = F(x; \alpha, \beta) = F\left(\frac{x}{\beta}; \alpha\right)$
  - The Chi-Squared Distribution
    - Let ν be a positive integer. Then a random variable X is said to have a chi-squared distribution with parameter ν if the pdf of X is the gamma density with $\alpha = \nu/2$ and $\beta = 2$. The pdf of a chi-squared rv is thus
      $$f(x; \nu) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \ge 0 \\ 0 & x < 0 \end{cases} \qquad (4.10)$$
      The parameter ν is called the number of degrees of freedom (df) of X. The symbol χ^2 is often used in place of "chi-squared."
- **4.5 Others**
  - 4.5 The Weibull Distribution
    - A random variable X is said to have a Weibull distribution with parameters α and β $(\alpha > 0, \beta > 0)$ if the pdf of X is
      $$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \ge 0 \\ 0 & x < 0 \end{cases} \qquad (4.11)$$
    - **Alpha = 1, Exponential. Beta = 1/lambda**
    - Alpha: Shape, Beta: Scale
    - $\mu = \beta\Gamma\left(1 + \frac{1}{\alpha}\right)$
    - $\sigma^2 = \beta^2\left\{\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right\}$
    - $F(x; \alpha, \beta) = \begin{cases} 0 & x < 0 \\ 1 - e^{-(x/\beta)^\alpha} & x \ge 0 \end{cases}$
  - 4.5 The Lognormal Distribution
    - A nonnegative rv X is said to have a lognormal distribution if the rv $Y = \ln(X)$ has a normal distribution. The resulting pdf of a lognormal rv when $\ln(X)$ is normally distributed with parameters μ and σ is
      $$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{-[\ln(x)-\mu]^2/(2\sigma^2)} & x \ge 0 \\ 0 & x < 0 \end{cases}$$
    - Mu and sigma regarindg Ln(X) not to X
    - $E(X) = e^{\mu + \sigma^2/2}$
    - $V(X) = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$
    - $F(x; \mu, \sigma) = P(X \le x) = P[\ln(X) \le \ln(x)] = P\left(Z \le \frac{\ln(x) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right) \quad x \ge 0$
  - 4.5 Beta Distribution
    - A random variable X is said to have a beta distribution with parameters α, β (both positive), A, and B if the pdf of X is
      $$f(x; \alpha, \beta, A, B) = \begin{cases} \frac{1}{B-A} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\frac{x-A}{B-A}\right)^{\alpha-1} \left(\frac{B-x}{B-A}\right)^{\beta-1} & A \le x \le B \\ 0 & \text{otherwise} \end{cases}$$
      The case $A = 0, B = 1$ gives the standard beta distribution.
    - $\mu = A + (B - A) \cdot \frac{\alpha}{\alpha + \beta}$
    - $\sigma^2 = \frac{(B-A)^2 \alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
- **4.6 Probability Plots**
  - An effective way to check a distributional assumption is to construct
  - Sample Percentiles
    - Order the n sample observations from smallest to largest. Then the ith smallest observation in the list is taken to be the $[100(i - .5)/n]$th sample percentile.
  - $\left(\begin{matrix}[100(i - .5)/n]\text{th percentile} \\ \text{of the distribution,}\end{matrix}, \begin{matrix}i\text{th smallest sample} \\ \text{observation}\end{matrix}\right)$
    - The plotted points will then fall close to a 45 degree line.
  - measurement error has a standard normal dis- tribution
  - Extreme
    - A plot of the n pairs
      $$([100(i - .5)/n]\text{th } z \text{ percentile, } i\text{th smallest observation})$$
      on a two-dimensional coordinate system is called a normal probability plot.
  - nonlinear probability axis.